



```
PR 05-APR-1999; 99US-0127706.
PR 21-APR-1999; 99US-0130359.
PR 27-APR-1999; 99US-0131270.
PR 27-APR-1999; 99US-0131272.
PR 27-APR-1999; 99US-0131291.
PR 04-MAY-1999; 99US-0132371.
PR 04-MAY-1999; 99US-0132379.
PR 04-MAY-1999; 99US-0132383.
PR 25-MAY-1999; 99US-0135750.
PR 08-JUN-1999; 99US-0138166.
PR 20-JUL-1999; 99US-0144791.
PR 03-AUG-1999; 99US-0146970.
PR 09-DEC-1999; 99US-0170262.
XX (GETH ) GENENTECH INC.
XX
XX Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
XX Stewart TA, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI: 2000-628263/60.
XX P-PSDB; AAB18927.
XX
XX Novel secreted and transmembrane polypeptides useful for diagnosing
XX tumour in a mammal, for identifying agonists and antagonists of the
XX polypeptide and for therapeutic use
XX
XX Claim 2; Fig 37; 222pp; English.
XX
XX The present sequence encodes a secreted or transmembrane polypeptide.
XX The specification describes polypeptides designated PRO1484, PRO4334,
XX PRO1122, PRO1889, PRO1890, PRO1887, PRO1785, PRO4353, PRO4357, PRO4405,
XX PRO4356, PRO4352, PRO4380, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990,
XX PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is
XX useful for diagnosing tumour in a mammal. The polypeptides, their
XX agonists and antagonists are useful treating a condition associated with
XX expression or activity of the polypeptide. Conditions treated include
XX obesity, diabetes or hyper- or hypo-insulinemia. The polypeptides are
XX capable of inducing proliferation of mammalian kidney mesangial cells
XX and are therefore useful for treating kidney disorders associated with
XX decreased mesangial cell function such as Bergers disease or other
XX nephropathies associated with Schonlein-Henoch purpura, celliac disease,
XX dermatitis herpetiformis or Crohns disease. The nucleic acids may be used
XX to generate transgenic animals for use in development and screening of
XX therapeutically useful reagents and also for chromosome identification
XX and tissue typing.
XX
XX Sequence 2379 BP; 574 A; 695 C; 604 G; 506 T; 0 other;

Query Match 89.9%; Score 98; DB 21; Length 2379;
Best Local Similarity 100.0%; Pred. No. 8e-17;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
Db 2277 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2336

QY 70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
Db 2337 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2374

RESULT 2
AAV21239
ID AAV21239 standard; cDNA; 144 BP.
XX
XX AAV21239;
XX
XX 03-AUG-1998 (first entry)
XX
XX Homo sapiens clone H698_3 3' end.
XX
XX secreted protein; human; PBMC; peripheral blood mononuclear cells; ds.
```

```
XX Homo sapiens.
XX WO9807859-A2.
XX
XX 26-FEB-1998.
XX
XX 22-AUG-1997; 97WO-US14874.
XX
XX 23-AUG-1996; 96US-0702344.
XX (GEMY ) Genetics Inst Inc.
XX
XX Merberg F, McCoy JM, Lavallie BR, Racie LA, Treacy M, Spaulding V;
XX Jacobs K;
XX WPI: 1998-169163/15.
XX
XX New nucleic acid encoding secreted proteins from human cells -
XX useful e.g. as immuno-modulators, antitumour agents, promoters of
XX tissue growth, haemostatic and thrombolytic agents etc.
XX
XX Claim 33; Page 62; 79pp; English.
XX
XX The sequence is that of the 3' end of clone H698_3 (including
XX the polyA tail) which was isolated from a human adult PBMC cDNA
XX library using methods selective for cDNAs that encode secreted
XX proteins.
XX
XX Sequence 144 BP; 132 A; 0 C; 0 G; 8 T; 4 other;

Query Match 89.0%; Score 97; DB 19; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
Db 20 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 79

QY 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
Db 80 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 116

RESULT 3
AAAX77331
ID AAAX77331 standard; DNA; 144 BP.
XX
XX AAAX77331;
XX
XX 04-AUG-1999 (first entry)
XX
XX Human secreted protein encoding DNA (clone H698-3) 3' portion.
XX
XX Secreted protein; retina; brain; blood; testis; nutritional activity;
XX cell proliferation; differentiation; immune stimulation; human;
XX hematopoiesis regulation; tissue growth; chemotactic; thrombolytic;
XX anti-inflammatory; tumour invasion suppression; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO9926973-A1.
XX
XX 03-JUN-1999.
XX
XX 20-NOV-1998; 98WO-US24944.
XX
XX 19-NOV-1998; 98US-0196027.
XX
XX 21-NOV-1997; 97US-0976112.
XX
XX (GEMY ) GENETICS INST INC.
```

PI Collins-Racie LA, Evans C, Jacobs K, Lavallie ER;  
 PI McCoy JM, Merberg D, Spaulding V, Treacy M;  
 XX  
 DR WPI; 1999-357814/30.  
 XX

PT New polynucleotides encoding secreted human proteins  
 PS  
 PS Disclosure; Page 97; 106pp; English.

XX The invention relates to human secreted proteins (AA21588-597) encoded  
 CC by polynucleotides obtained from adult retina, adult brain, adult blood  
 CC or adult testes cDNA libraries. The nucleic acid sequences encoding the  
 CC secreted proteins correspond to clones AV421-2, BV278-2, C544-1,  
 CC CC332-33, CC365-40, CG68-4, D329-1, H698-2 and H963-20, (all clones are  
 CC deposited as ATCC 98145). The PNs and proteins are predicted to have  
 CC biological activities which would make them suitable for treating,  
 CC preventing or ameliorating medical conditions in humans and animals,  
 CC although no supporting data is given. Suggested activities include  
 CC nutritional activity, cytokine and cell proliferation/differentiation  
 CC activity, immune stimulating (e.g. as vaccines) or suppressing activity,  
 CC hematopoiesis regulating activity, tissue growth activity, activin/  
 CC inhibin activity, chemotactic/chemokinetic activity, hemostatic and  
 CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumor invasion suppressor activity and tumor  
 CC inhibition activity. The PNs are also stated to be useful for gene  
 CC therapy.

XX Sequence 144 BP; 132 A; 0 C; 0 G; 8 T; 4 other;

Query Match 89.0%; Score 97; DB 20; Length 144;

Best Local Similarity 100.0%; Pred. No. 4.3e-16;

Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 tttaaaaaaa 89.0%; Score 97; DB 20; Length 144;

Db 20 tttaaaaaaa 100.0%; Pred. No. 4.3e-16;

QY 71 aaaaaaa 100.0%; Score 97; DB 20; Length 144;

Db 80 aaaaaaa 100.0%; Pred. No. 4.3e-16;

RESULT 4

AAH70047/c

ID AAH70047 standard; cDNA; 170 BP.

XX

AC AAH70047;

XX 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 1321.

DE Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

OS WO200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US33312.

XX 08-DEC-1999; 99US-0169681.

XX 21-DEC-1999; 99US-0171350.

XX 14-MAR-2000; 2000US-0189315.

XX 12-MAY-2000; 2000US-0203791.

XX 09-JUN-2000; 2000US-0210600.

XX 21-JUL-2000; 2000US-0220114.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

PI

XX WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer  
 PT and for assessing and detecting compounds for treating the cancer -

XX Claim 1; Page 309; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with  
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
 CC polypeptides are useful: to assess if a patient is afflicted with  
 CC cervical cancer or has a pre-malignant condition; to monitor the  
 CC progression of cervical cancer or a premalignant condition in a patient;  
 CC and to select and/or assess the efficacy of a compound or therapy for  
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
 CC useful for gene therapy.

XX Sequence 170 BP; 11 A; 11 C; 10 G; 137 T; 1 other;

Query Match 89.0%; Score 97; DB 22; Length 170;

Best Local Similarity 100.0%; Pred. No. 4e-16;

Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 tttaaaaaaa 89.0%; Score 97; DB 22; Length 170;

Db 163 TTTAAAAA 100.0%; Pred. No. 4e-16;

QY 71 aaaaaaa 100.0%; Score 97; DB 22; Length 170;

Db 103 AAAAAA 100.0%; Pred. No. 4e-16;

RESULT 5

AAAS01025

ID AAAS01025 standard; cDNA; 378 BP.

XX

AC AAAS01025;

XX 29-MAY-2001 (first entry)

XX Sugarcane plant gene promoter cDNA isolated from clone c32A.

XX Sugarcane promoter region; monocotyledonous plant; stem tissue;  
 KW insecticide; herbicide; disease resistance; improved food content;  
 KW beta-glucuronidase; GUS; starch biosynthesis; fatty acid biosynthesis;  
 KW ADP-glucose pyrophosphorylase; sucrose metabolism; clone c32A; ss.

OS Saccharum sp.

XX Key Location/Qualifiers

FT CDS 2..46

FT /\*tag= a

FT /partial

FT /note= "Amino acid sequence deduced from the

FT longest and conserved ORF. This sequence lacks

FT a start codon"

XX WO200118211-A1.

XX 15-MAR-2001.

XX 01-SEP-2000; 2000WO-AU01033.

XX 02-SEP-1999; 99AU-0002625.

XX (UYQU ) UNIV QUEENSLAND.

XX Potier B, Birch RG;

XX WPI; 2001-218560/22.

XX P-PSDB; AAU00452.

XX



CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
CC useful for gene therapy.

XX  
SQ Sequence 597 BP; 259 A; 54 C; 39 G; 244 T; 1 other;

Query Match 89.0%; Score 97; DB 22; Length 597;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaa 70  
|||||  
Db 141 TTTAA 82

Qy 71 aaa 107  
|||||  
Db 81 AA 45

RESULT 8  
AAH71551/C  
ID AAH71551 standard; cDNA; 621 BP.  
AC AAH71551;  
XX  
DT 19-SEP-2001 (first entry)  
XX  
DE Human cervical cancer marker nucleic acid 2825.  
XX  
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200142467-A2.  
XX  
PD 14-JUN-2001.  
XX  
PF 06-DEC-2000; 2000WO-US33312.  
XX  
PR 08-DEC-1999; 99US-0169681.  
PR 21-DEC-1999; 99US-0171350.  
PR 14-MAR-2000; 2000US-0189315.  
PR 12-MAY-2000; 2000US-0203791.  
PR 09-JUN-2000; 2000US-0210600.  
PR 21-JUL-2000; 2000US-0220114.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Deeds J, Berger A, Zhao X;  
XX  
DR WPI: 2001-375006/39.  
XX  
PT New isolated nucleic acid for diagnosing and treating cervical cancer  
PT and for assessing and detecting compounds for treating the cancer -  
XX  
PS Claim 1; Page 579; 1051pp; English.  
XX

CC The invention relates to novel genes (AAH68727-AAH73383) associated with  
CC cervical cancer with cytostatic activity. The nucleic acids and encoded  
CC polypeptides are useful: to assess if a patient is afflicted with  
CC cervical cancer or has a pre-malignant condition; to monitor the  
CC progression of cervical cancer or a premalignant condition in a patient;  
CC and to select and/or assess the efficacy of a compound or therapy for  
CC inhibiting cervical cancer in a patient. The nucleic acids may also be  
CC useful for gene therapy.  
XX  
SQ Sequence 621 BP; 273 A; 63 C; 47 G; 238 T; 0 other;

Query Match 89.0%; Score 97; DB 22; Length 621;  
Best Local Similarity 100.0%; Pred. No. 2.4e-16;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaa 70  
|||||  
Db 135 TTTAAA 76

Qy 71 aaa 107  
|||||  
Db 75 AA 39

RESULT 9  
AAAG4638  
ID AAAG4638 standard; DNA; 801 BP.  
XX  
AC AAAG4638;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Partial sequence MEL3 associated with melanoma and thyroid tumors.  
XX  
KW Melanoma; Thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;  
KW rectal cancer; lung cancer; breast cancer; colon cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200050595-A2.  
XX  
PD 31-AUG-2000.  
XX  
PF 25-FEB-2000; 2000WO-US04929.  
XX  
PR 25-FEB-1999; 99US-0257417.  
XX  
PA (GOUT/) GOUT I.  
PA (RODN/) RODNIN N.  
PA (FILO/) FILOENKO V.  
PA (MATS/) MATSUKA G.  
PA (SCAN/) SCANLAN M.  
PA (OLDL/) OLD L.  
PA (BILY/) BILYNSKY B.  
XX  
PI Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;  
PI Bilynsky B;  
XX  
DR WPI: 2000-572092/53.  
XX  
PT Novel isolated nucleic acid molecules for diagnosing and treating  
PT melanoma, thyroid tumors, rectal, lung, breast and colon cancers -  
XX  
PS Claim 1; Page 62; 94pp; English.

CC The specification describes polynucleotides which are associated with  
CC melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,  
CC Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for  
CC diagnosing and treating a patient with melanoma, thyroid tumour,  
CC rectal cancer, lung cancer, breast cancer or colon cancer. The  
CC present sequence represents a polynucleotide of the invention.  
XX  
SQ Sequence 801 BP; 178 A; 198 C; 162 G; 165 T; 98 other;

Query Match 89.0%; Score 97; DB 21; Length 801;  
Best Local Similarity 100.0%; Pred. No. 2.1e-16;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaa 70  
|||||  
Db 548 tttaa 707

Qy 71 aa 107  
|||||  
Db 708 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 744

```
RESULT 10
AA501021
ID AAS01021 standard; cDNA; 976 BP.
XX AC
XX AAS01021;
XX DT 29-MAY-2001 (first entry)
XX DE Sugarcane plant gene promoter cDNA isolated from clone c67.
XX KW Sugarcane promoter region; monocotyledonous plant; stem tissue;
XX insecticide; herbicide; disease resistance; improved food content;
XX beta-glucuronidase; GUS; starch biosynthesis; fatty acid biosynthesis;
XX ADP-glucose pyrophosphorylase; sucrose metabolism; clone c67; ss.
XX OS
XX Saccharum sp.
XX FH Key Location/Qualifiers
XX 5'UTR 1..32
XX FT /*tag= a
XX CDS 33..596
XX FT /*tag= b
XX FT /*note= "Amino acid sequence deduced from the
XX longest and conserved ORF"
XX 3'UTR 597..581
XX FT /*tag= c
XX polyA_site 882
XX FT /*tag= d
XX PN WO200118211-A1.
XX 15-MAR-2001.
XX PD
XX PF 01-SEP-2000; 2000WO-AU01033.
XX PR 02-SEP-1999; 99AU-0002625.
XX PA (UYOU ) UNIV QUEENSLAND.
XX PI Potier B, Birch RG;
XX WPI; 2001-218560/22.
XX P-PSDB; AAU000450.
XX PT New sugarcane plant promoters for directing expression of heterologous
XX nucleic acids in a constitutive or tissue-specific manner in
XX monocotyledonous plants -
XX PS Claim 1; Fig 14; 107pp; English.
XX CC The present sequence for sugarcane plant promoter cDNA isolated from
XX clone c67 is 1 of 11 promoter regions of a transcribable DNA sequence
XX isolated from various sugarcane cDNA clones (AAS01021-AA501031). Also
XX described are 4 promoter regions of specific transcribed DNA sequences
XX (AAS01032-AA501035). The nucleic acids are useful for producing
XX transgenic plants, having an altered phenotype and for driving expression
XX of a foreign or endogenous DNA sequence, which encode agronomic
XX properties including insecticide, herbicide, disease resistance,
XX stress tolerance and improved food content, or increased yields. The
XX foreign or endogenous DNA sequence may comprise a region transcribed
XX into an antisense RNA or ribozyme that modulates the expression of a
XX corresponding target gene, or it may encode beta-glucuronidase (GUS),
XX luciferase, neomycin phosphotransferase, a product conferring herbicide
XX tolerance, a product affecting starch biosynthesis or modification,
XX ADP-glucose pyrophosphorylase, a product involved in fatty acid
XX biosynthesis, a product conferring insect resistance, a product altering
XX sucrose metabolism or a gene encoding valuable pharmaceuticals, e.g.
XX antibiotics, secondary metabolites or vaccines. The promoters are capable
XX of directing high level expression in many or all cells of a plant,
XX preferentially in stem or meristem tissue of monocotyledonous plants.
XX Sequence 976 BP; 320 A; 213 C; 229 G; 214 T; 0 other;
```

```
Query Match 89.0%; Score 97; DB 22; Length 976;
Best Local Similarity 100.0%; Pred. No. 2e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
Db 879 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 938
Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
Db 939 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 975
RESULT 11
AAZ52527
ID AAZ52527 standard; cDNA; 1034 BP.
XX AC AAZ52527;
XX DT 29-FEB-2000 (first entry)
XX DE Human secreted protein clone ya66_1 nucleotide sequence SEQ ID NO:105.
XX KW Human; secreted protein; immunostimulatory; haemostatic; cytokine;
XX proliferative; differentiative; chemotactic; chemokinetic; vaccine;
XX thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
XX gene therapy; ss.
XX OS Homo sapiens.
XX PN WO9958642-A2.
XX PD 18-NOV-1999.
XX PF 14-MAY-1999; 99WO-US10843.
XX PR 14-MAY-1998; 98US-0085472.
XX PR 17-AUG-1998; 98US-0096824.
XX PR 11-SEP-1998; 98US-0099843.
XX PR 15-SEP-1998; 98US-0099950.
XX PR 29-SEP-1998; 98US-0100424.
XX PR 09-OCT-1998; 98US-0102329.
XX PR 11-DEC-1998; 98US-0103615.
XX PR 14-DEC-1998; 98US-0111799.
XX PR 31-DEC-1998; 98US-0114415.
XX PR 10-FEB-1999; 99US-0248059.
XX PR 06-APR-1999; 99US-0287150.
XX PR 13-MAY-1999; 99US-0311021.
XX PA (GEMY ) GENETICS INST INC.
XX PI Wong GG, Clark HF, Fecthel K, Agostino MJ;
XX WPI; 2000-053095/04.
XX P-PSDB; AAY73442.
XX PT Novel polynucleotides and proteins having biological activities which
XX make them suitable for treating, preventing or ameliorating medical
XX conditions in humans or animals -
XX PS Claim 114; Page 653; 730pp; English.
XX CC The present invention describes human secreted proteins encoded by
XX polynucleotides obtained from adult testes, foetal brain, adult brain,
XX brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
XX cDNA libraries. The polynucleotides and proteins are predicted to have
XX biological activities which would make them suitable for treating,
XX preventing or ameliorating medical conditions in humans and animals.
XX Suggested activities include nutritional activity, cytokine and cell
XX proliferation/differentiation activity, immune stimulating (e.g. as
XX vaccines) or suppressing activity, haematopoiesis regulating activity,
```



```
SQ Sequence 1164 BP; 420 A; 176 C; 195 G; 373 T; 0 other;

Query Match      89.0%; Score 97; DB 22; Length 1164;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
    |
Db 1039 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1098

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
    |
Db 1099 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1135

RESULT 14
ID AAX04325 standard; DNA; 1376 BP.
XX
AC AAX04325;
XX
DT 13-APR-1999 (first entry)
XX
DE Human secreted protein gene 15 clone HSPES04.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO956804-A1.
XX
PD 17-DEC-1998.
XX
PF 11-JUN-1998; 98WO-US12125.
XX
PR 02-OCT-1997; 97US-0061060.
PR 13-JUN-1997; 97US-0049547.
PR 13-JUN-1997; 97US-0049548.
PR 13-JUN-1997; 97US-0049549.
PR 13-JUN-1997; 97US-0049550.
PR 13-JUN-1997; 97US-0049606.
PR 13-JUN-1997; 97US-0049607.
PR 13-JUN-1997; 97US-0049608.
PR 13-JUN-1997; 97US-0049609.
PR 13-JUN-1997; 97US-0049610.
PR 13-JUN-1997; 97US-0049611.
PR 13-JUN-1997; 97US-0050566.
PR 13-JUN-1997; 97US-0050901.
PR 13-JUN-1997; 97US-0052989.
PR 08-JUL-1997; 97US-0051919.
PR 18-AUG-1997; 97US-0055984.
PR 12-SEP-1997; 97US-0058665.
PR 12-SEP-1997; 97US-0058668.
PR 12-SEP-1997; 97US-0058669.
PR 12-SEP-1997; 97US-0058750.
PR 12-SEP-1997; 97US-0058971.
PR 12-SEP-1997; 97US-0058972.
PR 12-SEP-1997; 97US-0058975.
PR 02-OCT-1997; 97US-0060834.
PR 02-OCT-1997; 97US-0060841.
PR 02-OCT-1997; 97US-0060844.
PR 02-OCT-1997; 97US-0060865.
PR 02-OCT-1997; 97US-0061059.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

SQ Sequence 1376 BP; 499 A; 231 C; 286 G; 359 T; 1 other;

Query Match      89.0%; Score 97; DB 20; Length 1376;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
    |
Db 1257 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1316

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
    |
Db 1317 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1353

RESULT 15
ID AAA27985 standard; cDNA; 1495 BP.
XX
AC AAA27985;
XX
DT 29-AUG-2000 (first entry)
XX
DE Corn CCR4 transcription factor nucleotide sequence #2.
XX
KW CCR4; transcription factor; transgenic plant; antibody production;
KW corn; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 1..1266
FT /tag= a
FT /product= "CCR4"
FT /note= "Transcription factor, no start codon given"
XX
WO200032783-A1.
XX
PD 08-JUN-2000.
XX
PF 01-DEC-1999; 99WO-US28334.
XX
```



PR 02-DEC-1998; 98US-0110585.  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
PA  
XX Allen SM, Sakai H;  
PI  
XX  
XX  
DR WPI: 2000-412329/35.  
DR P-PSDB; AAY94658.  
XX  
XX New nucleic acids encoding CCR4 transcription factor in plants and  
PT seeds, useful for producing transgenic plants, antibodies and selecting  
PT polynucleotides affecting transcription factor polypeptide expression  
PT  
XX  
PS Claim 2; Page 35-36; 49pp; English.  
XX  
XX This invention relates to nucleic acid fragments encoding CCR4  
CC transcription factors in plants and seeds. The present sequence  
CC represents a corn contig encoding CCR4. In saccharomyces cerevisiae the  
CC CCR4 gene encodes a transcription factor that is required for the  
CC regulation of a number of genes. CCR4 is a component of a multisubunit  
CC complex that activates expression of the glucose repressible alcohol  
CC dehydrogenase (ADH2) gene. The CCR4 gene suppresses ADH2 expression. The  
CC invention includes rice, corn, soybean and wheat CCR4 sequences. The  
CC polynucleotide sequences are useful for selecting an isolated  
CC polynucleotide that affects the expression levels of transcription factor  
CC polypeptide in a plant cell. This can be achieved by introducing the  
CC nucleotide sequence into a plant cell and comparing the level of CCR4  
CC polypeptide produced, with a plant cell that does not contain the  
CC polynucleotide. The polynucleotide sequences are useful as primers and  
CC probes for obtaining a nucleic acid fragment encoding a transcription  
CC factor polypeptide. An expression cassette comprising the CCR4  
CC polynucleotide sequence is useful for positive selection of a transformed  
CC cell by transforming it to a plant cell and growing the transformed plant  
CC cell and allowing expression of the polynucleotide in an amount  
CC sufficient to complement a CCR4 auxotroph to provide a positive selection  
CC means. Polypeptides encoded by the nucleotide sequences are useful for  
CC producing antibodies. The nucleotide sequences are also useful for  
CC producing transgenic plants that express altered levels of CCR4  
CC transcription factor, and for mapping the genes encoding CCR4  
XX transcription factor.  
SQ Sequence 1495 BP; 488 A; 314 C; 318 G; 375 T; 0 other;

Query Match 89.0%; Score 97; DB 21; Length 1495;  
Best Local Similarity 100.0%; Pred. No. 1.7e-16;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 tttaa 70  
Db 1378 ttttaa 70  
QY 71 aa 107  
Db 1438 aa 1474

Search completed: January 3, 2002, 10:27:59  
Job time: 3912 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:59:07 ; Search time 1166.04 Seconds  
(without alignments)  
1004.502 Million cell updates/sec

Title: US-09-594-596-1

Perfect score: 109

Sequence: 1 aaataaaatttaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO\_NUC

Gap 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:\*

1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_htc:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_htc:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	102	93.6	301	11	BF726451	BF726451 by06g04.y
2	102	93.6	405	10	AV682792	AV682792 AV682792
3	102	93.6	486	10	AV681638	AV681638 AV681638
4	102	93.6	513	10	AV755973	AV755973 AV755973
5	102	93.6	694	10	AV734386	AV734386 AV734386
6	101	92.7	434	10	AI954183	AI954183 w*80g05.x
7	101	92.7	562	10	AW636413	AW636413 b146d05.w
8	100	91.7	145	10	AU038865	AU038865 AU038865
9	100	91.7	187	10	AL079963	AL079963 DXF2p586L
10	100	91.7	383	10	AL045163	AL045163 DXF2p434A
11	100	91.7	876	12	AF116682	AF116682 Homo sapi
12	99	90.8	150	10	AU175731	AU175731 AU175731

13	99	90.8	164	10	AA117464	AA117464 mn22d09.r
14	99	90.8	180	10	AW598705	AW598705 sj94h09.y
15	99	90.8	245	10	AA638677	AA638677 VO56b09.r
16	99	90.8	267	10	AW087445	AW087445 XB98g05.x
c 17	99	90.8	334	10	AI537273	AI537273 tp07a04.x
18	99	90.8	343	10	BE057182	BE057182 sm97f08.y
19	99	90.8	354	10	AA575800	AA575800 vn63a02.r
20	99	90.8	360	10	BE330213	BE330213 so74g08.y
21	99	90.8	391	10	AV723745	AV723745 AV723745
22	99	90.8	394	10	AL042745	AL042745 DKF2p434C
23	99	90.8	476	10	AW096620	AW096620 EST289800
24	99	90.8	629	11	BF969228	BF969228 602489742
25	99	90.8	935	11	BE878014	BE878014 601489792
26	98	89.9	146	10	AA590017	AA590017 vm22e09.r
27	98	89.9	185	10	BE047852	BE047852 tz43b10.y
28	98	89.9	229	11	BI142829	BI142829 rk7aell.y
29	98	89.9	241	10	BE022702	BE022702 sm87d02.y
30	98	89.9	243	10	AA571665	AA571665 vm09e04.r
c 31	98	89.9	261	10	AW129106	AW129106 xe97d08.x
32	98	89.9	276	10	AI249323	AI249323 qx48a07.x
33	98	89.9	302	10	BE683281	BE683281 182759 MA
34	98	89.9	318	11	BE800962	BE800962 sr07d02.y
35	98	89.9	319	10	AW156548	AW156548 se28e08.y
36	98	89.9	368	10	BE022471	BE022471 sm74d03.y
37	98	89.9	375	10	AV759831	AV759831 AV759831
38	98	89.9	403	10	AV681858	AV681858 AV681858
39	98	89.9	417	10	BE538997	BE538997 601059713
40	98	89.9	439	11	BF727092	BF727092 by16f02.y
41	98	89.9	451	10	AV682195	AV682195 AV682195
42	98	89.9	515	11	BG622615	BG622615 602647359
43	98	89.9	535	10	BE749464	BE749464 200401 MA
44	98	89.9	562	11	BG285103	BG285103 602409377
45	98	89.9	562	11	BG621124	BG621124 602616814

#### ALIGNMENTS

RESULT 1  
BF726451 301 bp mRNA EST 05-JAN-2001  
LOCUS by06g04.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by06g04 5', mRNA sequence.  
DEFINITION sapiens cDNA clone by06g04 5', mRNA sequence.  
ACCESSION BF726451  
VERSION BF726451.1 GI:12042362  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 301)  
Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics  
JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
COMMENT Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 06 row: g column: 04  
Seq primer: M13Rpl reverse primer (ABI).  
Location/Qualifiers  
1. .301  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="by06g04"  
By "Human Lens cDNA (Un-normalized, unamplified): sapiens"  
/tissue\_type="Lens"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies, essentially following the protocols of the Superscript Plasmid System full details of which are contained in the manufacturer's instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTCAGATCGGAGCGGCCGC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

```
BASE COUNT      153 a    41 c    33 g    74 t
ORIGIN

Query Match      93.6%; Score 102; DB 11; Length 301;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaa 405 bp mRNA EST 25-SEP-2000
Db 10 AAAAAATTTAAAAA 405 bp mRNA EST 25-SEP-2000
Qy 66 aaaaaa 405 bp mRNA EST 25-SEP-2000
Db 70 AAAAAA 405 bp mRNA EST 25-SEP-2000

RESULT 2
LOCUS AV682792 405 bp mRNA EST 25-SEP-2000
DEFINITION AV682792 GKB Homo sapiens cDNA clone GKBAG07 5', mRNA sequence.
ACCESSION AV682792
VERSION AV682792.1 GI:10284655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
, Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
, Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
, Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens cDNA GK- clones
Unpublished (2000)
TITLE Chinese National Human Genome Center at Shanghai
JOURNAL 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
COMMENT Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
1..405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GKBAG07"
/clone_lib="GKB"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      195 a    57 c    61 g    92 t
ORIGIN

Query Match      93.6%; Score 102; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaa 405 bp mRNA EST 19-OCT-2000
Db 26 AAAAAATTTAAAAA 405 bp mRNA EST 19-OCT-2000
Qy 66 aaaaaa 405 bp mRNA EST 19-OCT-2000
Db 86 AAAAAA 405 bp mRNA EST 19-OCT-2000

RESULT 4
LOCUS AV755973 513 bp mRNA EST 19-OCT-2000
DEFINITION AV755973 BM Homo sapiens cDNA clone BMFAP03 5', mRNA sequence.
ACCESSION AV755973
VERSION AV755973.1 GI:10913821
KEYWORDS EST.
```

```
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaa 486 bp mRNA EST 25-SEP-2000
Db 33 AAAAAATTTAAAAA 486 bp mRNA EST 25-SEP-2000
Qy 66 aaaaaa 486 bp mRNA EST 25-SEP-2000
Db 93 AAAAAA 486 bp mRNA EST 25-SEP-2000

RESULT 3
LOCUS AV681638 486 bp mRNA EST 25-SEP-2000
DEFINITION AV681638 GKB Homo sapiens cDNA clone GKBAGH04 5', mRNA sequence.
ACCESSION AV681638
VERSION AV681638.1 GI:10283501
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,
, Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,
, Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang
, Y., Gu,Y., Chen,Z. and Han,Z.
Homo sapiens cDNA GK- clones
Unpublished (2000)
TITLE Chinese National Human Genome Center at Shanghai
JOURNAL 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
COMMENT Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
1..486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GKBAGH04"
/clone_lib="GKB"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      288 a    78 c    92 g    28 t
ORIGIN

Query Match      93.6%; Score 102; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaa 486 bp mRNA EST 19-OCT-2000
Db 26 AAAAAATTTAAAAA 486 bp mRNA EST 19-OCT-2000
Qy 66 aaaaaa 486 bp mRNA EST 19-OCT-2000
Db 86 AAAAAA 486 bp mRNA EST 19-OCT-2000

RESULT 4
LOCUS AV755973 513 bp mRNA EST 19-OCT-2000
DEFINITION AV755973 BM Homo sapiens cDNA clone BMFAP03 5', mRNA sequence.
ACCESSION AV755973
VERSION AV755973.1 GI:10913821
KEYWORDS EST.
```

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 513)  
 AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,  
 Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng  
 ,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,  
 Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.  
 TITLE Homo sapiens cDNA BM clones  
 JOURNAL  
 COMMENT Unpublished (2000)  
 CONTACT: zequang han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shouting Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919(ex.45)  
 Fax: 86-21-50801922  
 Email: hanzgchgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 source Location/Qualifiers  
 1..513  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="BMFAPE03"  
 /clone\_lib="BM"  
 /tissue\_type="Bone marrow"  
 /cell\_type="CD34+ hematopoietic stem/progenitor cell"  
 /lab\_host="BM25.8"  
 /note="vector: pTriplex2; Site\_1: sfIIA; Site\_2: sfIIB"  
 BASE COUNT 200 a 94 c 126 t 2 others  
 ORIGIN  
 Query Match 93.6%; Score 102; DB 10; Length 513;  
 Best Local Similarity 100.0%; Pred. NO. 1.5e-08;  
 Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 6 aaaaatttataaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65  
 Db 87 AAAAAATTAAAAAaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 146  
 Qy 66 aaaaaaiaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107  
 Db 147 AAAAAAiaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 188  
 RESULT 5  
 AV734386  
 LOCUS AV734386 cda Homo sapiens cDNA clone cdaAND06 5', mRNA sequence.  
 DEFINITION AV734386  
 ACCESSION AV734386  
 VERSION AV734386.1 GI:10851931  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 694)  
 AUTHORS Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N.,  
 Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu  
 ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J.,  
 Chen,Z. and Han,Z.  
 TITLE Homo sapiens cDNA cda clones  
 JOURNAL  
 COMMENT Unpublished (2000)  
 CONTACT: zequang han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shouting Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919(ex.45)  
 Fax: 86-21-50801922  
 Email: hanzgchgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 source Location/Qualifiers



DKFp586L1818 5', mRNA sequence.

ACCESSION AL079963  
VERSION AL079963.1 GI:5435539  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 187)

AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Koehrer, et al.)

JOURNAL Unpublished (1999)

COMMENT Contact: Koehrer K

MIPS

An Klöpferspitz 18a D-82152 Martinsried, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMF2 (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.

No sl sequence available.

This clone (DKFp586L1818) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES Location/Qualifiers

source

1..187

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DKFp586L1818"

/clone\_lib="586 (synonym: hutel)"

/tissue\_type="uterus"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI/MluI"

BASE COUNT 135 a 8 c 15 g 29 t

ORIGIN

Query Match 91.7%; Score 100; DB 10; Length 187;

Best Local Similarity 100.0%; Pred. No. 9.2e-08;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 aaatttaa 67

Db 77 AAATTAAAAAaaa 136

Qy 68 aa 107

Db 137 AA 176

RESULT 10

AL045163

LOCUS

DEFINITION DKFp434A1050\_r1 434 (synonym: htes3) Homo sapiens cDNA clone

ACCESSION DKFp434A1050 5', mRNA sequence.

VERSION AL045163

KEYWORDS EST.

SOURCE AL045163.2 GI:5935969

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 383)

AUTHORS Poustka, A., Klein, M., Mewes, H.W., Gassenhuber, J. and Wiemann, S.

TITLE EST (Poustka, et al.)

JOURNAL Unpublished (1999)

COMMENT On Jul 9, 1999 this sequence version replaced gi:5433326.

Contact: Poustka A.J.

Department Lehrach

Max-Planck-Institute for Molecular Genetics

Ilhnestrasse 73, 14195 Berlin, Germany

Tel: +49-30-84131623

Fax: +49-30-84131128

Email: poustka@mpg-berlin-dahlem.mpg.de

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKF2 (German Cancer Research Center,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone (DKFp434A1050) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..383

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="DKFp434A1050"

/clone\_lib="434 (synonym: htes3)"

/tissue\_type="testis"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Vector: pSport1; Site\_1: NotI; Site\_2: SalI"

BASE COUNT 164 a 66 c 85 t

ORIGIN 2 others

Query Match 91.7%; Score 100; DB 10; Length 383;

Best Local Similarity 100.0%; Pred. No. 4.2e-08;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 attcaa 69

Db 261 ATTTAAAAAaaa 320

Qy 70 aaagg 109

Db 321 AAAGG 360

RESULT 11

AF116682

LOCUS

DEFINITION Homo sapiens PRO2013 mRNA, complete cds.

ACCESSION AF116682

VERSION AF116682.1 GI:7959862

KEYWORDS human.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 876)

AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhou, G., Ouyang, S., Luo, L.,

Bi, J., Liu, M. and He, F.

TITLE Functional prediction of the coding sequences of 121 new genes

JOURNAL deduced by analysis of cDNA clones from human fetal liver

UNPUBLISHED

REFERENCE 2 (bases 1 to 876)

AUTHORS Zhang, C., Yu, Y., Zhang, S., Wei, H., Zhou, G., Ouyang, S., Luo, L.,

Bi, J., Liu, M. and He, F.

TITLE Direct Submission

JOURNAL Submitted (24-DEC-1998) Department of Experimental Hematology,

Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing

100850, P. R. China

FEATURES

source

1..876

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="FLB7547"

/tissue\_type="liver"

/dev\_stage="fetal"

136..381

/note="predicted protein of HQ2013"

CDS

/codon_start=1									
/product="PRO2013"									
/protein_id="AAF71102.1"									
/db_xref="GI:7959863"									
/translation="MNVDEHVNLLVEIHRLGSKNADGKLSVKFGVLFRDDKCANLFE"									
ALVGTIAAKRRKIVYPGELLIGQVHDDVDIILLQD"									
BASE COUNT	333 a	131 c	167 g	245 t					
ORIGIN									
Query Match 91.7%; Score 100; DB 12; Length 876;									
Best Local Similarity 100.0%; Pred. No. 1.7e-08; Indels 0; Gaps 0;									
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	8	aaatttaa	107	67					
Db	775	AAATTATAAA	834						
Qy	68	aaa	107						
Db	835	AA	874						
RESULT 12									
LOCUS	AU175731	150 bp	mRNA	EST	05-JUL-2001				
DEFINITION	AU175731 Marsupinaeus japonicus adult Marsupinaeus japonicus cDNA clone PJA376 reverse, mRNA sequence.								
ACCESSION	AU175731								
VERSION	AU175731.1 GI:14622138								
KEYWORDS	EST.								
SOURCE	Marsupinaeus japonicus.								
ORGANISM	Marsupinaeus japonicus								
REFERENCE	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea; Penaeidae; Marsupinaeus.								
AUTHORS	1 (bases 1 to 150)								
TITLE	Rojtinnakorn,J., Hirono,I., Aoki,T., Itami,T. and Takahashi,Y. Gene expression in non-infected and viral infected hemocytes of kuruma prawn (Penaeus japonicus)								
JOURNAL	Unpublished (2001)								
COMMENT	Contact: Takashi Aoki Aquatic Biosciences Tokyo University of Fisheries Konan 4-5-7, Tokyo, Minato-ku 108-8477, Japan Tel: 81-3-5463-0689 Fax: 81-3-5463-0690 Email: aoki@tokyo-u-fish.ac.jp.								
FEATURES	Location/Qualifiers								
source	1..150								
	/organism="Marsupinaeus japonicus"								
	/db_xref="taxon:27405"								
	/clone="PJA376"								
	/clone_lib="Marsupinaeus japonicus adult"								
	/cell_type="hemocytes"								
	/dev_stage="adult"								
	/note="Vector: lambda ZAP II"								
BASE COUNT	118 a	3 c	7 g	21 t	1 others				
ORIGIN									
Query Match 90.8%; Score 99; DB 10; Length 150;									
Best Local Similarity 100.0%; Pred. No. 1.7e-07; Indels 0; Gaps 0;									
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	9	aatttaa	68						
Db	43	AATTATAAA	102						
Qy	69	aaa	107						
Db	103	AA	141						

DEFINITION	sj94h09.y1 Cm-cl023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: 1	
ACCESSION	Gm-cl023-2298 5', mRNA sequence.	
VERSION	AW598705	
KEYWORDS	AW598705.1 GI:7286218	
SOURCE	EST.	
ORGANISM	soybean.	
	Glycine max	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	
REFERENCE	1. (bases 1 to 180)	
AUTHORS	Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Stepec,M., Theising,B., Allen,M., Bowers,Y., Peterson,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.	
TITLE	Public Soybean EST Project	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Shoemaker R/Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: estewartson.wustl.edu This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones.genomesystems.com or info.genomesystems.com web site: www.genomesystems.com Putative full length read vector to vector length is.	

FEATURES	source
Location/Qualifiers	
1. .180	
/organism="Glycine max"	
/db_xref="taxon:3847"	
/clone="GENOME SYSTEMS CLONE ID: Gm-c1023-2298"	
/clone_lib="Gm-c1023"	
/tissue_type="seed coats of greenhouse grown plants"	
/lab_host="DHI08"	
/note="Vector: pSPORtl; Site_1: SalI; Site_2: NotI; This cDNA library was constructed from mRNA isolated from seed coats (100-200mgs) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restrictions site. Sal I linkers adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORtl vector. The ligated cDNA fragments were transformed into E.coli Electromax DHI08 host cells. This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."	
114 a	5 c 20 g 41 t
BASE COUNT	
ORIGIN	

		90.8%;	Score 99;	DB 10;	Length 180;
	Best Local Similarity	100.0%;	Pred. No. 1.4e-07;		
	Matches 99;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	9	aatttaa	68		
Db	81	AATTTAA	140		
Qy	69	aaa	107		
Db	141	AAA	179		

Search completed: January 3, 2002, 10:25:00  
Job time: 5153 sec

RESULT 15	AA638677	245 bp	mRNA	EST	22-OCT-1997
LOCUS	AA638677	245 bp	mRNA	EST	22-OCT-1997
DEFINITION	AA638677	245 bp	mRNA	EST	22-OCT-1997
ACCESSION	AA638677	245 bp	mRNA	EST	22-OCT-1997
VERSION	AA638677	245 bp	mRNA	EST	22-OCT-1997
KEYWORDS	AA638677	245 bp	mRNA	EST	22-OCT-1997
SOURCE	AA638677	245 bp	mRNA	EST	22-OCT-1997
ORGANISM	AA638677	245 bp	mRNA	EST	22-OCT-1997
REFERENCE	AA638677	245 bp	mRNA	EST	22-OCT-1997
AUTHORS	AA638677	245 bp	mRNA	EST	22-OCT-1997
TITLE	AA638677	245 bp	mRNA	EST	22-OCT-1997
JOURNAL	AA638677	245 bp	mRNA	EST	22-OCT-1997
COMMENT	AA638677	245 bp	mRNA	EST	22-OCT-1997

```

FEATURES
source
1. 245
Location/Qualifiers
/organism="Mus musculus"
/strain="FVE/N"
/db_xref="taxon:10090"
/clone="IMAGE:1053881"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT73D-PAC (Pharmacia) with a modified
polylinker; Site 1: EcoRI; Site 2: NotI; Tissue obtained
from 8 week old mouse. Colon was harvested 72 hours after
irradiation with 1400 Gys. 1st strand cDNA was primed
with a Not I oligo(dT) primer
5'-TGTTGAATCTGAGTGGAGGCGGCCCTTTTCTTTTCTTTTCTTTT
T 3'; double-stranded cDNA was ligated to Eco RI
adaptors [AATTCGGATCCGTG], digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library constructed by Bob Barstead."
150 a 28 c 24 g 43 t
BASE COUNT

```

```

BASE COUNT      150 a      28 c      24 g      43 t
ORIGIN

Query Match      90.8%;   Score 99;   DB 10;   Length 245;
Best Local Similarity 100.0%;   Pred. NO. 9.8e-08;
Matches 99;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```

[illegible]

Search completed: January 3, 2002, 10:25:00  
Job time: 5153 sec



---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:57:37 ; Search time 1342.2 seconds  
(without alignments)  
1339.735 Million cell updates/sec

Title: US-09-594-596-1  
Perfect score: 109  
Sequence: 1 aaataaaatttaaaaa.....aaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Result No.	Score	Query			DB ID	Description
		Match	Length	DB		
1	101	92.7	951	9	HSM801406	AL133560 Homo sapi
2	101	92.7	1843	9	BC009470	BC009470 Homo sapi
3	101	92.7	2506	9	AK026927	AK026927 Homo sapi
4	100	91.7	1827	9	AB070106	AB070106 Macaca fa
5	98	89.9	437	3	AF155369	AF155369 Butus ma
6	98	89.9	1370	9	AB060826	AB060826 Macaca fa
7	98	89.9	2295	9	AF090903	AF090903 Homo sapi
8	97	89.0	144	6	I89947	I89947 Sequence 26
c	97	89.0	173	6	AX185626	AX185626 Sequence
9	97	89.0	311	3	AF159977	AF159977 Butus ma
10	97	89.0	396	6	AX093192	AX093192 Sequence
11	97	89.0	597	6	AX187052	AX187052 Sequence
12	97	89.0	617	8	AF1323605	AF1323605 Manihot e
13	97	89.0	621	6	AX187131	AX187131 Sequence
c	97	89.0	641	8	AF266463	AF266463 Manihot e
15	97	89.0	868	8	AF049923	AF049923 Petunia x
16	97	89.0	1117	6	AR123145	AR123145 Sequence
17	97	89.0	1117	8	AF243375	AF243375 Glycine m
18	97	89.0	1164	6	AX098194	AX098194 Sequence
19	97	89.0	1489	9	AB048953	AB048953 Macaca fa
20	97	89.0	1851	9	AK026744	AK026744 Homo sapi
21	97	89.0	1906	9	AK025414	AK025414 Homo sapi
22	97	89.0	1926	9	HSM802868	AL512733 Homo sapi
23	97	89.0	1962	6	AX098509	AX098509 Sequence
24	97	89.0	2013	10	BC006019	BC006019 Mus muscu
25	97	89.0	2032	9	AK027204	AK027204 Homo sapi
26	97	89.0	2463	10	BC005449	BC005449 Mus muscu
27	97	89.0	2844	9	HSM802647	AL353940 Homo sapi
28	97	89.0	2943	10	BC005526	BC005526 Mus muscu
29	97	89.0	310	6	A08913	A08913 H sapiens (
30	96	88.1	310	6	AX187085	AX187085 Sequence
c	96	88.1	353	11	CNS06EJ3	AL393173 T7 end of
31	96	88.1	363	3	AF135821	AF135821 Mesobuthu
32	96	88.1	363	3	AF155365	AF155365 Butus ma
33	96	88.1	384	3	AF155365	AX093215 Sequence
34	96	88.1	396	6	AX093215	Y11790 A.thaliana
c	96	88.1	440	8	ATP25A	AL137550 Homo sapi
36	96	88.1	750	9	HSM802288	BC003501 Mus muscu
37	96	88.1	801	10	BC005501	Z83268 B.floridae
38	96	88.1	944	3	BFZ83268	X68215 P.sativum m
39	96	88.1	1022	8	PSPIAA45R	X82434 H.sapiens m
40	96	88.1	1252	9	HSSTA	AF083228 Caenorhab
41	96	88.1	1412	3	AF083228	A65341 Sequence 64
42	96	88.1	1474	6	A65341	AR150479 Sequence
43	96	88.1	1474	6	AR150479	AB063070 Macaca fa
44	96	88.1	1555	9	AB063070	
45	96	88.1				

ALIGNMENTS

RESULT 1

HSM801406	HSM801406	951 bp	mRNA	PRI	18-FEB-2000
LOCUS	Homo sapiens mRNA; CDNA DKFZp434M1414 (from clone DKFZp434M1414); partial cds.				
DEFINITION	AL133560				
ACCESSION	AL133560.1	GI:6599130			
VERSION					
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 951)				
AUTHORS	Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.				
TITLE	Submitted (15-DEC-1999); MIPS, Am Klopferspitz 18a, D-82152				
JOURNAL	Martinried, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by BWFZ (Biomedical Research Center at the Charite,				

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DRFzp434M1414) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: [clone@rzpd.de](mailto:clone@rzpd.de) Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES	Location/Qualifiers
source	1..951
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="DKF2p43M1414"
	/clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"
	/dev_stage="adult"
	/tissue_type="testis"
gene	1..513
CDS	/gene="DKF2p43M1414"
	<1..513
	/note="KIAA0829, partly"
	/codon_start=1
	/product="hypothetical protein"
	/protein_id="CAB63714.1"
	/db_xref="GI:6599131"
	/translation="ELIIRVEMGPFKHTVDDGLDIRKAAPFCMTLLDSCLDRLDIFIE FLNHVEDGKLDHYDIRMLTFLMLVRLSTLCPASVQLRDLRVLPLRATCTTKVKANS KQEFKQDELKRSAMRAVAALLTIPEAKSKPLMSFQSISSNPGLAAIFESIQKSDS SNLESMDIS"
polyA_signal	836..841
polyA_site	853
BASE COUNT	365 a 134 c 185 g 267 t
ORIGIN	
Query Match	92.7%; Score 101; DB 9; Length 951;
Best Local Similarity	100.0%; Pred. No. 1.4e-29;
Matches 101; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	7 aaatttttaa 66 
Db	847 AATTTTAA 906 
QY	67 aa 107 
Db	907 AA 947 
RESULT	2
BC009470	
LOCUS	BC009470 1843 bp mRNA PRI 12-JUL-2001
DEFINITION	Homo sapiens, protein kinase, interferon-inducible double stranded RNA dependent activator, clone MGC:1978 IMAGE:3503747, mRNA, complete cds.
ACCESSION	BC009470
VERSION	BC009470.1 GI:14495716
KEYWORDS	MGC.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Strausberg R.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabbs-remail.nih.gov">cgabbs-remail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Ruben Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
info@bcgsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smallos, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasya van den Bosch, Jill Vardy,  
George Yang, Scott Zuiderveld, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 6 Row: d Column: 22  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA qi: 3290197.

FEATURES source

source	CDS
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

CDS

[illegible]Query Match  
Best Local  
Matches 1

Qy	7 aa
D <sub>b</sub>	1717 AA

Qy 67 aa  
pb 1777 AA

RESULT 3  
AK026927

LOCUS	DEFINITION	ACCESSION
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		

VERSION  
KEYWORDS  
SOURCE

[illegible]

## AUTHORS

Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.  
 NEDO human cDNA sequencing project  
 JOURNAL Unpublished (2000)  
 REFERENCE 2 (bases 1 to 2506)  
 AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
 Shibahara,T., Tanaka,T. and Nakamura,Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio  
 Sugano, Institute of Medical Science, University of Tokyo,  
 Laboratory of Genome Structure Analysis, Human Genome Center,  
 Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan  
 [E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,  
 Fax:81-3-5449-5416]  
 COMMENT NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing; Research Association for Biotechnology; cDNA library  
 construction, 5'- & 3'-end one pass sequencing; Department of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).

## FEATURES

Location/Qualifiers  
 source

1..2506  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="HepG2"  
 /cell\_type="hepatoma"  
 /clone="HEP02623"  
 /clone\_lib="HEP"  
 /note="cloning vector pME18SFL3"

misc\_feature 1..2506  
 /note="highly similar to AF161354 Homo sapiens HSPC091  
 mRNA"

BASE COUNT 839 a 450 c 474 g 743 t  
 ORIGIN

Query Match 92.7%; Score 101; DB 9; Length 2506;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-29;  
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 aaaaatttaaaaaa 106  
 Db 2406 AAAAAA 2465  
 QY 66 aaaaaa 106  
 Db 2466 AAAAAA 2506

RESULT 4  
 AB070106 1827 bp mRNA PRI 16-AUG-2001  
 LOCUS Macaca fascicularis testis cDNA clone:QtsA-14022, full insert  
 DEFINITION sequence.  
 AB070106  
 ACCESSION AB070106.1 GI:15208052  
 VERSION  
 KEYWORDS Oligo capping; fis (full insert sequence).  
 SOURCE Macaca fascicularis adult male testis cDNA to mRNA,  
 clone\_lib:macaque testis cDNA library QtsA clone:QtsA-14022.  
 ORGANISM Macaca fascicularis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 1827)  
 AUTHORS Hashimoto,K., Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirai,M.,  
 Terao,K. and Sugano,S.  
 TITLE Isolation of novel full-length cDNA clones from macaque testis cDNA  
 JOURNAL libraries  
 REFERENCE 2 (bases 1 to 1827)  
 AUTHORS Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-AUG-2001) Katsuyuki Hashimoto, National Institute of

Infectious Diseases, Division of Genetic Resources; 23-1, Toyama  
 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan  
 (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,  
 Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)  
 Lab host: TOP10  
 Vector: pME18S-FL3 (Acc.No. AB009864)  
 R. Site1: DraIII (CAGCTGTG)  
 R. Site2: DraIII (CAGCTGTG)  
 Description: 1st strand cDNA was primed with an oligo(dT) primer  
 [ATGTGGCTTTT] and double-stranded cDNA was synthesized  
 using specific 5' and 3' primers and amplified by PCR. The PCR  
 product was digested with SfiI and size selection was performed to  
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned  
 into distinct DraIII sites of pME18S-FL3. XhoI sites just outside  
 the DraIII sites can be used to isolate the cDNA insert. Libraries  
 were constructed by oligo-capping method (Sugano et al., University  
 of Tokyo, Institute of Medical  
 Science). Custom primer used for sequencing  
 ( 5' end primer [CTCTGCTCTAAAGCTGG];  
 3' end primer [CGACTGCAGCTCGACACA] ).

## FEATURES

Location/Qualifiers  
 source

1..1827  
 /organism="Macaca fascicularis"  
 /db\_xref="taxon:9541"  
 /clone="QtsA-14022"  
 /sex="male"  
 /tissue\_type="testis"  
 /clone\_lib="macaque testis cDNA library QtsA"  
 /dev\_stage="adult"  
 132..1691  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="BA963051.1"  
 /db\_xref="GI:15208053"  
 /translation="MSEMSIDRNHNLSPGIPISVOTESWLDLILGKVSRRKKTAS  
 KVKAEPCQLASKLREKVINPEPKLNILIELEFKEDFTLPEPSLRTLPISGPPPI  
 LAYKESSNLDINFKDEEETSPKCFEGSDLRAFFSNVDCSEKSHGCCAFQNL  
 IDYIYBEQIKSPKALIAIDPHAAHYSEVRLKAKAKALQKQORMARHPAISR  
 EOTHFEEDDSKRLKTTISYOLVDIPEKQITDIDIVDFOLRNSNMSITCCDSRIACGKV  
 VRNLELKHVHGSKELTSPDGTTOIFYPSGNLAIIRVPKNVNGFTCIVOEDMPTNP  
 AILVALDSSGSSCYHPNGVWVYINLGGQYSDQAGNRIRANWNSINISSPASFVK  
 PFLALNRYIGVRLIFODKISITFLAMGQQAIRSVGTVKLPNPEIPVLURYVSGDDL  
 LLASITKIRLRFHKLKGCYVFPSSQVWEKLPKOPSYLSSLTLKLIALCRSSGKQDTM  
 KTRININEBI"

## CDS

BASE COUNT 702 a 350 c 332 g 443 t  
 ORIGIN

Query Match 91.7%; Score 100; DB 9; Length 1827;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-29;  
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 aaatttaaaaaa 107  
 Db 1684 AAATTTAAAA 1743  
 QY 68 aaaaaa 107  
 Db 1744 AAAAAA 1783

## RESULT 5

AF155369  
 LOCUS AF155369 437 bp mRNA INV 13-JUN-2001  
 DEFINITION Buthus martensii putative potassium ion channel blocker TKXsl mRNA,  
 complete cds.  
 ACCESSION AF155369  
 VERSION AF155369.1 GI:14388602  
 KEYWORDS Chinese scorpion.  
 SOURCE Buthus martensii  
 ORGANISM Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;  
 Buthoidea; Buthidae; Buthus.



```
Query Match      89.9%; Score 98; DB 9; Length 2395;
Best Local Similarity 100.0%; Pred. No. 1.9e-28;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 attcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
    |||||||
Db 2261 ATTTAAAAA 2320
    |||||||

Qy 70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa.107
    |||||||
Db 2321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2358
    |||||||

RESULT 8
LOCUS      I89947      144 bp      DNA      PAT      10-AUG-1998
DEFINITION Sequence 26 from patent US 5723315.
ACCESSION  I89947
VERSION     I89947.1 GI:3409887
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 144)
AUTHORS     Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M. and Spaulding,V.
TITLE        Secreted proteins and polynucleotides encoding them
JOURNAL      Patent: US 5723315-A 26 03-MAR-1998;
FEATURES     Location/Qualifiers
            source
            1..144
            /organism="unknown"
BASE COUNT   132 a 0 c 0 g 8 t 4 others
ORIGIN

Query Match      89.0%; Score 97; DB 6; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
    |||||||
Db 20 TTTAAAAA 79
    |||||||

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
    |||||||
Db 80 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 116
    |||||||

RESULT 9
LOCUS      AX185626/c 173 bp      DNA      PAT      06-AUG-2001
DEFINITION Sequence 1321 from Patent WO0142467.
ACCESSION  AX185626
VERSION     AX185626.1 GI:15137029
KEYWORDS   human.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 173)
AUTHORS     Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE        Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
JOURNAL      Patent: WO 0142467-A 1321 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES     Location/Qualifiers
            source
            1..173
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT   14 a 11 c 137 t 1 others
ORIGIN
```

```
Query Match      89.0%; Score 97; DB 6; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
    |||||||
Db 163 TTTAAAAA 104
    |||||||

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
    |||||||
Db 103 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 67
    |||||||

RESULT 10
LOCUS      AF159977      311 bp      mRNA      INV      13-JUN-2001
DEFINITION Buthus martensii toxin TLP3 mRNA, partial cds.
ACCESSION  AF159977
VERSION     AF159977.1 GI:14388612
KEYWORDS   .
SOURCE      Chinese scorpion.
ORGANISM    Buthus martensii
REFERENCE   1 (bases 1 to 311)
AUTHORS     Shunyi,Z. and Wenxin,L.
TITLE        cDNA encoding a toxin-like peptide from Bmk
JOURNAL      Unpublished
FEATURES     Location/Qualifiers
            source
            1..311
            /organism="Buthus martensii"
            /db_xref="taxon:34649"
            /tissue_type="venom gland"
            <1..142
            /codon_start=2
            /product="toxin TLP3"
            /protein_id="AAK61824.1"
            /db_xref="GI:14388613"
            /translation="ENLGEDCNLKQOKATDGCROPHCFTDMPDNYATRPDVTDP
            IM"
BASE COUNT   181 a 29 c 33 g 68 t
ORIGIN

Query Match      89.0%; Score 97; DB 3; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
    |||||||
Db 208 TTTAAAAA 267
    |||||||

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
    |||||||
Db 268 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 304
    |||||||

RESULT 11
LOCUS      AX093192/c 396 bp      DNA      PAT      30-MAR-2001
DEFINITION Sequence 10 from Patent WO0118046.
ACCESSION  AX093192
VERSION     AX093192.1 GI:13509641
KEYWORDS   human.
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```



Best Local Similarity 100.0%; Pred. No. 4.5e-28;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 tttaa 70  
|||||  
Db 135 TTTAAAAA  
QY 71 aa 107  
|||||  
Db 75 AAAAAA

RESULT 15  
AF266463  
LOCUS AF266463 641 bp mRNA PLN 24-MAY-2001  
DEFINITION Manihot esculenta RNA polymerase subunit mRNA, complete cds.  
ACCESSION AF266463  
VERSION AF266463.1 GI:14193246  
KEYWORDS  
SOURCE cassava.  
ORGANISM Manihot esculenta  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
Malpighiales; Euphorbiaceae; Manihot.  
1 (bases 1 to 641)  
REFERENCE  
AUTHORS Reilly, K., Winter, M.J., Han, Y., Tohme, J. and Beeching, J.R.  
TITLE RNA polymerase subunit isolated from cassava storage roots  
JOURNAL Unpublished  
2 (bases 1 to 641)  
AUTHORS Reilly, K., Winter, M.J., Han, Y., Tohme, J. and Beeching, J.R.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-2000) Biology and Biochemistry, University of  
Bath, Claverton Down, Bath BA2 7AY, UK  
FEATURES  
sourse 1..641  
/organism="Manihot esculenta"  
/db\_xref="taxon:3983"  
69..359  
/codon\_start=1  
/product="RNA polymerase subunit"  
/protein\_id="AAK55847.1"  
/db\_xref="GI:14193247"  
/translation="MVGDKFTMALAHTLNLDGTPDTGYIIGARKTLADKYEYIMHG  
KLYKISEGSGKAVKAEIYVSYGGLMLMKGDPSHSHFELDQRLFLIRKL"

BASE COUNT 257 a 92 c 116 g 176 t  
ORIGIN

Query Match 89.0%; Score 97; DB 8; Length 641;  
Best Local Similarity 100.0%; Pred. No. 4.5e-28;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 tttaa 70  
|||||  
Db 543 TTTAAAAA  
QY 71 aa 107  
|||||  
Db 603 AAAAAA

Search completed: January 3, 2002, 10:05:11  
Job time: 4054 sec









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 1164.61 Seconds  
(without alignments)  
1005.736 Million cell updates/sec

Title: US-09-594-596-1  
Perfect score: 109  
Sequence: 1 aaataaaatttaaaaaa.....aaaaaaaggg 109

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estro:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: gb\_estl1:\*  
11: gb\_estl2:\*  
12: gb\_hic:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pin:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105.4	96.7	301	11	BF726451
2	105.4	96.7	405	10	AV682792
3	105.4	96.7	486	10	AV681638
4	105.4	96.7	513	10	AV755973
5	105	96.3	576	13	CNS02568
6	104.6	96.0	340	10	AL513755
7	104.6	96.0	363	10	AL514473
8	104.2	95.6	200	10	AW185559
9	104.2	95.6	449	10	AL514879
10	104.2	95.6	1042	11	BG115502
11	104.2	95.6	1435	11	BF984813
12	103.8	95.2	134	13	CNS03F8P

13	103.8	95.2	142	10	AA117240
14	103.8	95.2	147	10	AI340603
15	103.8	95.2	267	10	AW087445
16	103.8	95.2	386	10	AV759054
17	103.8	95.2	402	10	AV682849
18	103.8	95.2	415	10	BE689881
19	103.8	95.2	449	10	AL513999
20	103.8	95.2	504	10	AV758806
21	103.8	95.2	542	10	AV757028
22	103.8	95.2	578	10	AW706506
23	103.8	95.2	683	10	AV682476
24	103.8	95.2	694	10	AV734386
25	103.8	95.2	712	10	AV734318
26	103.8	95.2	735	10	AV681769
27	103.8	95.2	762	10	AV758110
28	103.8	95.2	773	10	AA528822
29	103.8	95.2	803	11	BG335466
30	103.8	95.2	838	10	AL514075
31	103.8	95.2	935	11	BE878014
32	103.8	95.2	1016	13	CNS00359
33	103.8	95.2	1166	11	BE964682
34	103.4	94.9	400	10	AL513741
35	103.4	94.9	626	10	AL513937
36	103.2	94.7	147	10	AV610565
37	103.2	94.7	204	11	BI024203
38	103.2	94.7	273	10	AW164794
39	103	94.5	956	13	CNS014V0
40	102.8	94.3	120	10	AI284517
41	102.8	94.3	264	10	AA607075
42	102.8	94.3	278	10	AA638755
43	102.8	94.3	329	10	AL514627
44	102.8	94.3	383	10	AL045163
45	102.8	94.3	408	10	AL514793

## ALIGNMENTS

RESULT 1  
BF726451 301 bp mRNA EST 05-JAN-2001  
LOCUS by06g04.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by06g04 5', mRNA sequence.  
ACCESSION BF726451  
VERSION BF726451.1 GI:12042362  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.  
NIBANK: EST analysis and bioinformatics for ocular genomics  
Invest. Ophthalmol. Vis. Sci. 41 (2000) In press  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 06 row: g column: 04  
Seq primer: M13RP1 reverse primer (ABI).  
Location/Qualifiers  
1. .301  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="by06g04"  
/clone\_lib="Human Lens cDNA (Un-normalized, unamplified): BY:  
/tissue\_type="Lens"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"

```
/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses
from different adults (both approximately 40 years old)
together yielded 20ug of total RNA and 150ng mRNA for cDNA
library synthesis. A directionally cloned cDNA library in
the pCMVSPORT6 vector was constructed at Life Technologies
, essentially following the protocols of the Superscript
Plasmid System full details of which are contained in the
manufacturer's Instruction manual
(http://www.lifetech.com/). First strand synthesis was
carried out using a Not I primer-adaptor
[5'-pGACTAGTTCGATCGGAGCGCCGCTT)15-3']. Not I/blunt
end inserts were cloned into the Not I/EcoR V sites in the
vector. EST analysis was performed on the unamplified
library at the NIH Intramural Sequencing Center (NISC)."
```

BASE COUNT 153 a 41 c 33 g 74 t

Query Match 96.7%; Score 105.4; DB 11; Length 301;  
Best Local Similarity 99.1%; Pred. No. 3.8;  
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 aaataaaatttataaa 486 bp mRNA EST 25-SEP-2000  
||||| 60  
Db 5 AAAAAAAAAATTTAA 64  
Qy 61 aaaaaa 107  
Db 65 AA 111

RESULT 2  
AV682792 405 bp mRNA EST 25-SEP-2000  
LOCUS AV682792 GKB Homo sapiens cDNA clone GKBA07 5', mRNA sequence.  
DEFINITION AV682792  
ACCESSION AV682792  
VERSION AV682792.1 GI:10284655  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 405)  
AUTHORS Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,  
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,  
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang  
,Y., Gu,Y., Chen,Z. and Han,Z.  
Homo sapiens cDNA GK- clones  
Unpublished (2000)  
TITLE  
JOURNAL  
COMMENT Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzgchgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..405  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GKBAG07"  
/clone\_lib="GKB"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
Location/Qualifiers  
1..405  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GKBAG07"  
/clone\_lib="GKB"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
XhoI"

BASE COUNT 195 a 57 c 61 g 92 t

Query Match 96.7%; Score 105.4; DB 10; Length 405;

Best Local Similarity 99.1%; Pred. No. 3.1;  
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 aaataaaatttataaa 486 bp mRNA EST 25-SEP-2000  
||||| 60  
Db 28 AAAAAAAAAATTTAA 87  
Qy 61 aaaaaa 107  
Db 88 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 134

RESULT 3  
AV681638 486 bp mRNA EST 25-SEP-2000  
LOCUS AV681638 GKB Homo sapiens cDNA clone GKBAJH04 5', mRNA sequence.  
DEFINITION AV681638  
ACCESSION AV681638  
VERSION AV681638.1 GI:10283501  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 486)  
AUTHORS Wu,T., Qian,B., Huang,Q., Kang,B., Gao,X., Xu,Z., Xiao,H.,  
Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L.,  
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang  
,Y., Gu,Y., Chen,Z. and Han,Z.  
Homo sapiens cDNA GK- clones  
Unpublished (2000)  
TITLE  
JOURNAL  
COMMENT Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzgchgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..486  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GKBAJH04"  
/clone\_lib="GKB"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

BASE COUNT 288 a 78 c 92 g 28 t

Query Match 96.7%; Score 105.4; DB 10; Length 486;  
Best Local Similarity 99.1%; Pred. No. 2.7;  
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 aaataaaatttataaa 486 bp mRNA EST 25-SEP-2000  
||||| 80  
Db 21 AAAAAAAAAATTTAA 80  
Qy 61 aaaaaa 107  
Db 81 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 127

RESULT 4  
AV755973 513 bp mRNA EST 19-OCT-2000  
LOCUS AV755973 BM Homo sapiens cDNA clone BMFAPE03 5', mRNA sequence.  
DEFINITION AV755973  
ACCESSION AV755973  
VERSION AV755973.1 GI:10913821  
KEYWORDS EST.





```

RESULT 9
AL514879/c AL514879 449 bp mRNA EST 13-FEB-2001
DEFINITION AL514879 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BB014ZE01 3
prime, mRNA sequence.
ACCESSION AL514879
VERSION AL514879.1 GI:12778372
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 449)
AUTHORS Li.W.B., Gruber.C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
Location/Qualifiers
1..449
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CL0BB014ZE01"
/tissue_type="placenta"
/note="vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 96 a 107 c 19 g 194 t 33 others
ORIGIN

Query Match 95.6%; Score 104.2; DB 10; Length 449;
Best Local Similarity 96.3%; Pred. No. 3.5;
Matches 103; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 aaataaaatttataaaataaaataaaataaaataaaataaaataaaataaa 60
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 120 AAATAAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAA 61

QY 61 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 107
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 60 AAAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 14

RESULT 10
BG115502
LOCUS BG115502 1042 bp mRNA EST 30-JAN-2001
DEFINITION BG115502 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4416974 5',
mRNA sequence.
ACCESSION BG115502
VERSION BG115502.1 GI:12609008
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1042)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC

```

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10148 row: n column: 15
High quality sequence stop: 391.
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4416974"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 446 a 219 c 265 g 112 t
ORIGIN

Query Match 95.6%; Score 104.2; DB 11; Length 1042;
Best Local Similarity 97.2%; Pred. No. 1.9;
Matches 106; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 aaataaaatttataaaataaaataaaataaaataaaataaaataaaataaa 60
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 253 AAATAAACATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAA 312

QY 61 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 109
|||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 313 AAAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 361

RESULT 11
BF984813/c BF984813 1435 bp mRNA EST 23-JAN-2001
DEFINITION BF984813 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4399310 5',
mRNA sequence.
ACCESSION BF984813
VERSION BF984813.1 GI:12387625
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1435)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10102 row: n column: 15
High quality sequence start: 8
High quality sequence stop: 133.
Location/Qualifiers
1..1435
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4399310"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"

```



/lab_host="DH10B (phage-resistant)" /site="Organ: small intestine; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."									
BASE COUNT	589 a	164 c	500 g	182 t					
ORIGIN									
Query Match	95.6%	Score 104.2;	DB 11;	Length 1435;					
Best Local Similarity	97.2%	Pred. No. 1.5;	3;	Indels 0;	Gaps 0;				
Matches 106;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;					
Qy	1	aaataaaattt	aaa	60					
Db	134	AAATAAAAAAT	ATAACAAAAA	AAA	75				
Qy	61	aaaaaaaaaaaa	aaa	109					
Db	74	AAAAAAAAAAAA	AAA	AGG 26					
RESULT 12									
CNS03F8P	134 bp	DNA	GSS	17-MAY-2000					
LOCUS									
DEFINITION	Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone 021P15 of library G from Tetraodon nigroviridis, genomic survey sequence.								
ACCESSION	AL241378								
VERSION	AL241378.1 GI:7962147								
KEYWORDS	GSS; genome survey sequence.								
SOURCE	Tetraodon nigroviridis.								
ORGANISM	Tetraodon nigroviridis								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.								
AUTHORS	1 (bases 1 to 134) Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billaud,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.								
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis								
JOURNAL	Unpublished								
REFERENCE	2 (bases 1 to 134) Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brotier,P., Quetier,F., Saurin,W. and Weissenbach,J.								
AUTHORS	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence								
TITLE	Unpublished								
JOURNAL	3 (bases 1 to 134) Genoscope.								
REFERENCE	Direct Submission								
AUTHORS	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases								
TITLE	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at								
COMMENT	http://www.genoscope.cns.fr/Tetraodon. Location/Qualifiers								
FEATURES	1..134 /organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="021P15" /clone_lib="G" /note="Genoscope sequence ID : COBG021CH08SP1-end : PUC-Ori"								
BASE COUNT	120 a	2 c	1 g	10 t	1 others				
ORIGIN									
Query Match	95.2%	Score 103.8;	DB 13;	Length 134;					

[illegible]

Search completed: January 3, 2002, 09:42:30  
Job time: 2733 sec

RESULT	LOCUS	AW087445	267 bp	mRNA	EST	15-OCT-1999
	DEFINITION	xh98rc05 x1 NCI CGAP Co21 Homo sapiens cDNA clone IMAGE:358260 v1				

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 100.12 Seconds  
(without alignments)  
933.364 Million cell updates/sec

Title: US-09-594-596-1

Perfect score: 109

Sequence: 1 aaataaaatttaaaaa.....aaaaaaaaaaaaaaagg 109

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103.8	95.2	1860	22	AAAF72806
2	103.4	94.9	3060	22	AAAD02924
3	103.4	94.9	3060	22	AAAF44263
4	103.4	94.9	3060	22	AAAC91489
5	102.6	94.1	225	21	AAAC98734
6	102.6	94.1	396	22	AAAF94842
7	102.6	94.1	464	22	AAH33241
8	102.6	94.1	569	21	AAH37052
9	102.6	94.1	612	21	AAAC99862
10	102.6	94.1	938	22	AAAD07722
11	102.6	94.1	1048	21	AAA26373

c

12	102.6	94.1	1051	21	AAZ91918	Murine mahogany pr
13	102.6	94.1	1126	21	AAAC59409	Human secreted pro
14	102.6	94.1	1164	22	AAAF98699	Human secreted pro
15	102.6	94.1	1254	21	AAAC98626	Human ovarian can
16	102.6	94.1	1492	21	AAAC98102	Human secreted pro
17	102.6	94.1	1493	22	AAH34433	Human colon cancer
18	102.6	94.1	1510	21	AAAC9961	Human colon cancer
19	102.6	94.1	1733	20	AAV81394	Murine TANGO 183 c
20	102.6	94.1	1868	22	AAAD08428	Human tumour anti
21	102.6	94.1	1868	22	AAAD08428	Human secreted pro
22	102.6	94.1	1868	22	AAAD08428	NANBH virus strain
23	102.2	93.8	671	22	AAAD05318	Human secreted pro
24	102.2	93.8	903	21	AAAC79727	Human secreted pro
25	102.2	93.8	2772	22	AAAF92074	Human PRO1277 cDNA
26	102.2	93.8	2773	21	AAZ65009	Membrane-bound pro
27	102.2	93.8	2773	21	AAZ65009	Membrane-bound pro
28	101.6	93.2	160	22	AAAF44155	Human PRO1277 (UNO
29	101.6	93.2	160	22	AAAF44155	Retroviral recombi
30	101.6	93.2	174	21	AAAC98740	Human colon cancer
31	101.6	93.2	178	22	AAAF24361	Retroviral recombi
32	101.6	93.2	208	21	AAAC98557	Human colon cancer
33	101.6	93.2	208	22	AAAF24363	Retroviral recombi
34	101.6	93.2	712	22	AAH34312	Human colon cancer
35	101.6	93.2	1708	22	AAAF26570	DNA encoding human
36	101.6	93.2	2240	21	AAAC93391	Human secreted pro
37	101.6	93.2	2240	22	AAAD08375	Human secreted pro
38	101.6	93.2	2240	22	AAAD08375	Human secreted pro
39	101.6	93.2	2240	22	AAAD08375	Human secreted pro
40	101.6	93.2	2240	22	AAAD08375	Human secreted pro
41	101.6	93.2	2240	22	AAAD08375	Human secreted pro
42	101.6	93.2	2240	22	AAAD08375	Human secreted pro
43	101.6	93.2	2240	22	AAAD08375	Human secreted pro
44	101.6	93.2	2240	22	AAAD08375	Human secreted pro
45	101.6	93.2	2240	22	AAAD08375	Human secreted pro

## ALIGNMENTS

RESULT	1
AAAF72806	AAAF72806 standard; DNA; 1860 BP.
ID	AAAF72806 standard; DNA; 1860 BP.
AC	AAAF72806;
DT	24-APR-2001 (first entry)
DE	Secreted protein gene #8.
KW	Secreted protein; human; autoimmune; hyperproliferation; cardiovascular; cerebrovascular; infection; food; ds.
OS	Homo sapiens.
PN	WO200107459-A1.
PD	01-FEB-2001.
PF	20-JUL-2000; 2000WO-US19735.
PR	23-JUL-1999; 99US-0145220.
XX	(HUMA-) HUMAN GENOME SCI INC.
XX	Rosen CA, Ruben SM, Ebner R, Duan RD, Ni J, Soppet DR, Moore PA;
XX	Shi Y, Lafleur DW, Olsen HS, Komatsoulis GA;
XX	WPI; 2001-123261/13.
XX	New isolated nucleic acid encoding 29 secreted proteins, for
XX	diagnosing, preventing and treating e.g. autoimmune,
XX	hyperproliferative, cardiovascular, and ocular diseases or disorders
XX	and microorganism infections

```
XX
PS
XX
CC The present invention relates to 29 human secreted proteins. The
CC invention is used to prevent autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the
CC breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections
CC caused by bacteria, viruses and fungi and ocular disorders e.g.
CC corneal infection. Also used in food preparations.
XX
XX Sequence 1860 BP; 588 A; 370 C; 410 G; 490 T; 2 other;

Query Match          95.2%; Score 103.8; DB 22; Length 1860;
Best Local Similarity 98.1%; Pred. No. 0.00032;
Matches 105; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 aaataaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
Db 1737 aaaaaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1796

QY 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
Db 1797 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1843

RESULT 2
AAD02924
ID AAD02924 standard; cDNA; 3060 BP.
AC AAD02924;
XX
XX 31-MAY-2001 (first entry)
XX
XX Human PRO5723 cDNA (DNA82361).
XX
KW Human; PRO5723; antiinflammatory; dermatological; immunosuppressive;
KW antirheumatic; antiarthritis; osteopathic; antianemic; haemostatic;
KW antihypertensive; antidiabetic; antiviral; antipsoriatic; antiallergic;
KW antiasthmatic; inhibitor; therapy; systemic lupus erythematosus;
KW spondyloarthritis; systemic sclerosis; systemic vasculitis;
KW sarcoidosis; idiopathic inflammatory myopathy; Sjogren's syndrome;
KW autoimmune thrombocytopenia; immune-mediated renal disease; hepatitis;
KW demyelinating polyneuropathy; Guillain-Barre syndrome; Whipple's disease;
KW hepatobiliary disease; primary biliary cirrhosis; sclerosing cholangitis;
KW inflammatory bowel disease; gluten-sensitive enteropathy; skin disease;
KW allergic rhinitis; atopic dermatitis; food hypersensitivity; urticaria;
KW eosinophilic pneumonia; hypersensitivity pneumonitis; graft rejection;
KW idiopathic pulmonary fibrosis; graft-versus-host-disease; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH 77..1135
FT CDS /*tag= a
FT /*product= "Human PRO5723 protein"
FT sig_peptide 77..133
FT /*tag= b
FT mat_peptide 134..1132
FT /*tag= c
FT /*product= "Mature human PRO5723 protein"
XX
XX WO200116319-A2.
XX
XX 08-MAR-2001.
PD
XX
XX 23-AUG-2000; 2000WO-US23522.
PF
XX
XX 31-AUG-1999; 99US-0151733.
PR 01-SEP-1999; 99WO-US20111.
PR 16-DEC-1999; 99WO-US30095.
PR
```

```
PR 18-FEB-2000; 2000WO-US04342.
PR 01-MAR-2000; 2000WO-US05601.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 05-JUN-2000; 2000US-0209832.
XX (GETH ) GENENTECH INC.
PA
XX Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Tumas D;
PI Watanabe CK, Wood WI;
XX
XX WPI; 2001-226690/23.
DR P-PSDB; AAY72878.
XX
XX New PRO polypeptides for treating immune related and inflammatory
XX diseases such as rheumatoid arthritis, systemic vasculitis, asthma,
XX autoimmune hemolytic anemia, and diabetes mellitus .
PT
PT Claim 2; Fig 9; 118pp; English.
XX
XX The present sequence is a cDNA (DNA82361 clone) encoding PRO5723
XX protein. PRO protein, its agonist or antagonist or its antibody which
XX are capable of enhancing or inhibiting the proliferation of T-lymphocytes
XX or of increasing the infiltration of inflammatory cells into a tissue
XX are useful in the diagnosis and treatment of immune-related diseases
XX in mammals. The PRO protein is useful for treating systemic lupus
XX erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic
XX arthritis, spondyloarthritis, systemic sclerosis, idiopathic
XX inflammatory myopathy, Sjogren's syndrome, systemic vasculitis,
XX sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia,
XX thyroiditis, diabetes mellitus, immune-mediated renal disease.
XX demyelinating disease of the central or peripheral nervous system,
XX idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
XX chronic inflammatory demyelinating polyneuropathy, hepatitis, primary
XX disease, infectious or autoimmune chronic active hepatitis, primary
XX biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
XX inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
XX disease, autoimmune or immune-mediated skin diseases such as bullous
XX skin disease, erythema multiforme and contact dermatitis, psoriasis,
XX allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
XX food hypersensitivity and urticaria, immunologic diseases of the lung
XX such as eosinophilic pneumonias, idiopathic pulmonary fibrosis, hyper-
XX sensitivity pneumonitis, transplantation associated diseases such as
XX graft rejection or graft-versus-host-disease.
XX
XX Sequence 3060 BP; 1094 A; 559 C; 569 G; 837 T; 1 other;
SO
```

```
Query Match          94.9%; Score 103.4; DB 22; Length 3060;
Best Local Similarity 97.2%; Pred. No. 0.00029;
Matches 104; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 aaataaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
Db 2929 aaataaaagttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2988

QY 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
Db 2989 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3035

RESULT 3
AAF44263
ID AAF44263 standard; cDNA; 3060 BP.
XX
AC AAF44263;
XX
XX 02-APR-2001 (first entry)
XX
XX Human PRO5723 nucleotide sequence SEQ ID NO:504.
DE
XX
```

Human, secreted and transmembrane protein; PRO; cytostatic;  
cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
diagnostic assay; ss.

Homo sapiens.

WO200073454-A1.

07-DEC-2000.

30-MAR-2000; 2000WO-US08439.

02-JUN-1999; 99WO-US12252.

23-JUN-1999; 99US-0141037.

07-JUL-1999; 99US-0143048.

20-JUL-1999; 99US-0144758.

26-JUL-1999; 99US-0145698.

28-JUL-1999; 99US-0146222.

17-AUG-1999; 99US-0149396.

15-SEP-1999; 99WO-US21090.

15-SEP-1999; 99WO-US21547.

08-OCT-1999; 99US-0158663.

30-NOV-1999; 99WO-US28313.

01-DEC-1999; 99WO-US28301.

16-DEC-1999; 99WO-US30095.

20-DEC-1999; 99WO-US30911.

05-JAN-2000; 2000WO-US00219.

06-JAN-2000; 2000WO-US00376.

11-FEB-2000; 2000WO-US03565.

18-FEB-2000; 2000WO-US04341.

22-FEB-2000; 2000WO-US04414.

24-FEB-2000; 2000WO-US04914.

24-FEB-2000; 2000WO-US05004.

03-MAR-2000; 2000WO-US05841.

15-MAR-2000; 2000WO-US06884.

20-MAR-2000; 2000WO-US07377.

(GETH ) GENENTECH INC.

Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
Grimaldi CJ, Gurney AL, Klavin IJ, Napier MA, Pan J, Paoni NF;  
Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
Zhang Z;

WPI; 2001-032160/04.

P-PSDB; AAB65294.

PRO polynucleotides used to produce polypeptides used to target  
bioactive molecules such as toxins, radiolabels or antibodies, to  
specific cells, to cause targeted cell death -

Claim 2; Fig 317; 935pp; English.

The present invention describes human secreted and transmembrane PRO  
proteins. The PRO proteins have cytostatic activity. The PRO proteins  
can be used for targeted delivery of bioactive molecules, such as  
toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
sequences, and their fragments, can be used as hybridisation probes, in  
chromosomal and gene mapping, and in the generation of anti-sense RNA  
and DNA. They may also be used to produce transgenic animals which are  
used to develop and screen therapeutically useful reagents. The PRO  
nucleotide and protein sequence can be used for tissue typing and in  
treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
AA44270 to AA44470 represent PCR primers and hybridisation probes used  
in the isolation of human PRO sequences. AA44087 to AA44269 and  
AA65154 to AAB65300 represent human PRO polynucleotide and protein  
sequences given in the exemplification of the present invention.

Sequence 3060 BP; 1094 A; 559 C; 569 G; 837 T; 1 other;

Query Match

94.9%; Score 103.4; DB 22; Length 3060;

Best Local Similarity 97.2%; Pred. NO. 0.00029;  
Matches 104; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 aaataataatttataaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60  
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

DB 2929 aaataaagtttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2988  
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

OY 61 aa 107  
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

DB 2989 aa 3035  
||||||| ||||||||| ||||||||| ||||||||| ||||||||| ||||||||| |||||||||

RESULT 4

AAC91489

ID AAC91489 standard; cDNA; 3060 BP.

XX AAC91489;

DT 21-MAR-2001 (first entry)

XX Human PRO5723 cDNA.

XX Human; PRO; antiinflammatory; dermatological; antiarthritic;

KW antiinflammatory; cardiac; antianaemic; immunosuppressive; antithyroid;

KW antidiabetic; nootropic; neuroprotective; hepatotropic; virucide;

KW antiallergic; antiasthmatic; immune related disorder;

KW hepatobiliary disease; autoimmune disease; allergy; ss.

XX Homo sapiens.

OS WO200073452-A2.

PN 07-DEC-2000.

XX 02-JUN-2000; 2000WO-US15264.

XX 02-JUN-1999; 99WO-US12252.

PR 20-JUL-1999; 99US-0144732.

PR 20-JUL-1999; 99US-0144758.

PR 28-JUL-1999; 99US-0146222.

PR 01-SEP-1999; 99WO-US20111.

PR 15-SEP-1999; 99WO-US21090.

PR 29-SEP-1999; 99US-0162506.

PR 30-NOV-1999; 99WO-US28313.

PR 01-DEC-1999; 99WO-US28634.

PR 09-DEC-1999; 99US-0170262.

PR 20-DEC-1999; 99WO-US30911.

PR 03-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 18-FEB-2000; 2000WO-US04342.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 15-MAR-2000; 2000WO-US06884.

PR 20-MAR-2000; 2000WO-US07377.

PR 21-MAR-2000; 2000WO-US07532.

PR 30-MAR-2000; 2000WO-US08439.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Chan B, Goddard A, Godowski PJ, Gurney AL;

PI Hebert C, Henzel W, Kabakoff RC, Shelton DL, Tumas D, Watanabe CK;

PI Wood WI;

XX WPI; 2001-025253/03.

DR P-PSDB; AAB50930.

XX

XX

PT

PT

PT

PT

PT

PT

PT

PT systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
XX thyroiditis and diabetes mellitus -  
PS Claim 48; Fig 57; 218pp; English.  
XX  
CC The present sequence is one of thirty three nucleic acids encoding PRO  
CC polypeptides. The PRO polypeptides, anti-PRO antibodies, agonists and  
CC antagonists are useful for treating and diagnosing immune related  
CC disorders such as systemic lupus erythematosus, rheumatoid arthritis,  
CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,  
CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's  
CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic  
CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,  
CC immune-mediated renal disease, demyelinating diseases of the central  
CC and peripheral nervous systems (such as multiple sclerosis, idiopathic  
CC demyelinating polynuropathy or Guillain-Barre syndrome, and chronic  
CC inflammatory demyelinating polynuropathy), hepatobiliary diseases  
CC (such as infectious, autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis and sclerosing cholangitis),  
CC inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's  
CC disease, autoimmune or immune-mediated skin diseases (such as bullous  
CC skin diseases, erythema multiforme, contact dermatitis, psoriasis),  
CC allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,  
CC food hypersensitivity and urticaria), immunological diseases of the  
CC lung (such as eosinophilic pneumonias, idiopathic pulmonary fibrosis  
CC and hypersensitivity pneumonitis), transplantation associated diseases  
CC including graft rejection and graft-versus-host diseases.  
XX  
SQ Sequence 3060 BP; 1094 A; 559 C; 569 G; 837 T; 1 other;

Query Match 94.9%; Score 103.4; DB 22; Length 3060;  
Best Local Similarity 97.2%; Pred. NO. 0.00029;  
Matches 104; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 aaataaaatttataaaaaaa 60  
Db 2929 aaataaaagttaaaaaaa 2988  
Qy 61 aaaaaaa 107  
Db 2989 aaaaaaa 3035

RESULT 5  
AAC98734  
ID AAC98734 standard; cDNA; 225 BP.  
XX  
AC AAC98734;  
XX  
DT 09-MAR-2001 (first entry)  
XX  
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:744.  
XX  
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;  
KW neural disorder; immune system disorder; muscular disorder;  
KW reproductive disorder; gastrointestinal disorder; renal disorder;  
KW infectious disease; cardiovascular disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO20005351-A1.  
XX  
XX 21-SEP-2000.  
XX  
XX 08-MAR-2000; 2000WO-US05883.  
XX  
XX 12-MAR-1999; 99US-0124270.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;  
XX  
DR WPI: 2000-587534/55.  
DR P-PSDB; AAB53977.  
XX  
PT Colon cancer associated gene sequences, referred to as colon cancer  
PT antigens, useful for the treatment, prevention, and diagnosis of colon  
PT disorders such as colon cancer -  
XX  
XX Claim 1; Page 1304; 2104pp; English.  
XX  
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The  
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;  
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
CC vulnary, nephrotropic, antiinfective and antibacterial activities, and  
CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
CC proteins and antibodies to the proteins are useful for the prevention,  
CC treatment and diagnosis of colon disorders, such as colon cancer. The  
CC polynucleotides may be used in diagnostics and research, such as for  
CC chromosome identification, and as hybridisation probes. The proteins  
CC may also be used to prevent diseases such as neural disorders, immune  
CC system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, wounds, renal disorders, infectious  
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
CC AAB54007 represent sequences used in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 225 BP; 161 A; 11 C; 27 G; 20 T; 6 other;

Query Match 94.1%; Score 102.6; DB 21; Length 225;  
Best Local Similarity 96.3%; Pred. NO. 0.00094;  
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 aaataaaatttataaaaaaa 60  
Db 97 aaaaaaa 156  
Qy 61 aaaaaaa 109  
Db 157 aaaaaaa 205

RESULT 6  
AAF94842/c  
ID AAF94842 standard; cDNA; 396 BP.  
XX  
AC AAF94842;  
XX  
DT 23-MAY-2001 (first entry)  
XX  
DE Human ovarian cancer associated coding sequence SEQ ID NO: 33.  
XX  
KW Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200118046-A2.  
XX  
XX 15-MAR-2001.  
XX  
XX 08-SEP-2000; 2000WO-US24827.  
XX  
XX 10-SEP-1999; 99US-0394374.  
XX  
XX 01-MAY-2000; 2000US-0561778.  
XX  
XX 15-AUG-2000; 2000US-0640173.  
XX  
XX 07-SEP-2000; 2000US-0656668.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Xu J, Stolk JA;  
XX  
PI





PT useful for diagnosis and treatment of e.g. cancers, neurological  
PT disorders, immune diseases, inflammation or blood disorders -  
XX  
PS Claim 1; Page 337; 586pp; English.

XX AA927019 to AA927137 represent 94 isolated human secreted protein genes.  
CC AA986215 to AA986333 are the secreted proteins encoded by the 94 human  
CC genes. This sequence represents a fragment of one of the human secreted  
CC proteins. The genes and their corresponding secreted polypeptides are  
CC useful for preventing, treating or ameliorating medical conditions,  
CC e.g., by protein or gene therapy. Also pathological conditions can be  
CC diagnosed by determining the amount of the new polypeptides in a sample  
CC or by determining the presence of mutations in the new genes. Specific  
CC uses are described for each of the 94 genes, based on which tissues they  
CC are most highly expressed in, and include developing products for the  
CC diagnosis or treatment of cancer, tumours, developmental abnormalities  
CC and foetal deficiencies, blood disorders, diseases of the immune system,  
CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive  
CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin  
CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney  
CC disorders, digestive/endocrine disorders, infections and AIDS. The  
CC polypeptides are also useful for identifying their binding partners.  
CC The sequences shown in AA986334 to AA986585 represent fragments of the  
CC secreted proteins.

XX Sequence 569 BP; 204 A; 133 C; 137 G; 95 T; 0 other;

Query Match 94.1%; Score 102.6; DB 21; Length 569;  
Best Local Similarity 96.3%; Pred. No. 0.00066;  
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aaataaaatttataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 60  
Db 453 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 512

Qy 61 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 109

Db 513 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 561

RESULT 9  
AAC99862  
ID AAC99862 standard; cDNA: 612 BP.

XX AAC99862;

XX 13-MAR-2001 (first entry)

XX Human secreted protein gene 45 SEQ ID NO:55.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;  
KW fungicide; ophthalmological; gene therapy; pathological condition;  
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
KW neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;  
KW cerebrovascular disorder; angiogenesis; nervous system disorder;  
KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
KW wound healing; skin aging; food additive; preservative; ss.

XX Homo sapiens.

OS WO200070042-A1.

PN 23-NOV-2000.

XX 11-MAY-2000; 2000WO-US12788.

XX 13-MAY-1999; 99US-0134068.

XX (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;  
PI Duan RD, Florence KA, Soppet DR;  
XX WPI: 2000-679828/66.  
DR P-PSDB: AAB56121.

XX Isolated nucleic acid molecule encoding a human secreted protein is  
PT used in preventing, treating or ameliorating a medical condition -

PS Claim 1; Page 881; 1065pp; English.

XX The polynucleotide sequences given in AAC99818 to AAC99977 encode the  
CC human secreted proteins given in AAB56077 to AAB56362. Human secreted  
CC proteins have activities based on the tissues and cells the genes are  
CC expressed in. Examples of activities include: immunosuppressive; cardiant;  
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
CC vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;  
CC virucide; fungicide; and ophthalmological. The human secreted  
CC polynucleotides and proteins can be used to prevent, treat or ameliorate  
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
CC dogs, chickens or sheep. They are also used in diagnosing a pathological  
CC condition or susceptibility to a pathological condition. Disorders which  
CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid  
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or  
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular  
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders  
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
CC fungi and ocular disorders e.g. corneal infection. The proteins can also  
CC be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The proteins can also be used as a  
CC food additive or preservative to increase or decrease storage  
CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used  
CC in the exemplification of the present invention.

XX Sequence 612 BP; 199 A; 139 C; 139 G; 135 T; 0 other;

Query Match 94.1%; Score 102.6; DB 21; Length 612;  
Best Local Similarity 96.3%; Pred. No. 0.00064;  
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aaataaaatttataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 60

Db 496 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 555

Qy 61 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 109

Db 556 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 604

RESULT 10  
AAD07722  
ID AAD07722 standard; cDNA: 938 BP.

XX AAD07722;

XX 10-AUG-2001 (first entry)

XX Human secreted protein-encoding gene 18 cDNA clone HLYDU43, SEQ ID NO:28.

XX Human; secreted protein; proliferative disorder; cancer; tumour;  
KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;  
KW cardiovascular disorder; angiogenic disorder; kidney disorder;  
KW gastrointestinal disorder; pregnancy-related disorder;  
KW endocrine disorder; infection; wound healing; vulnery;  
KW cell culture; chemotaxis; food additive; gene therapy;  
KW binding partner identification; chromosome 2; ss.

```
XX OS Homo sapiens.
XX FH Location/Qualifiers
XX CDS 61..387
FT /product= "Human secreted protein precursor"
FT /tag= a
FT sig_peptide 61..138
FT mat_peptide 139..384
FT /tag= b
FT /tag= c
FT /product= "Human mature secreted protein"
XX WO200134800-A1.
XX 17-MAY-2001.
XX 08-NOV-2000; 2000WO-US30674.
XX 12-NOV-1999; 98US-0164750.
XX 30-JUN-2000; 2000US-0215128.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Ruben SM, Komatsoulis GA, Ebner R, Fiscella M, Wei P;
XX WPI; 2001-329085/34.
XX P-PSDB; AAEO3309.
XX New nucleic acid molecules encoding human secreted proteins, used in
XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
XX Parkinson's diseases and cancers -
XX Claim 1; Page 427; 530pp; English.
XX
XX AAD07705-AAD07759 represent cDNAs corresponding to 19 human secreted
XX protein genes, and AAEO3292-AAEO3346 represent the proteins they encode.
XX AAEO3347-AAEO3375 represent human secreted protein fragments or variants.
XX The genes and their secreted proteins are useful for preventing,
XX treating or ameliorating medical conditions, e.g., by protein or gene
XX therapy. Pathological conditions can be diagnosed by determining the
XX amount of the new protein in a sample or by determining the presence of
XX mutations in the new genes. Specific uses are described for each of the
XX 19 genes, based on the tissues in which they are most highly expressed,
XX and include developing products for the diagnosis or treatment of
XX proliferative disorders, cancer, tumours, foetal and developmental
XX abnormalities, haematopoietic disorders, diseases of the immune system,
XX AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
XX allergies, neurological disorders (e.g., Alzheimer's disease,
XX Parkinson's disease), cognitive disorders, schizophrenia, asthma,
XX skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, pregnancy-related disorders, endocrine
XX disorders, and infections. The proteins can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues, to identify their
XX cognate ligands or binding partners, and in chemotaxis, and can be used
XX as a food additive or preservative to modify storage properties.
XX Antibodies specific for a protein of the invention can be used in
XX alleviating symptoms associated with the disorders mentioned above, and
XX in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
XX immunosorbent assay (ELISA). The present sequence represents a human
XX secreted protein-encoding cDNA of the invention.
XX
XX Sequence 938 BP; 388 A; 143 C; 152 G; 253 T; 2 other;
XX
XX Query Match 94.1%; Score 102.6; DB 22; Length 938;
XX Best Local Similarity 96.3%; Pred. No. 0.00055;
XX Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 1 aaataaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
```

```
Db 793 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 852
Qy 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 109
Db 853 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 901

RESULT 11
AAA26373
ID AAA26373 standard; cDNA; 1048 BP.
XX
XX AAA26373;
XX 29-JUN-2000 (first entry)
XX Human secreted protein gene 28 SEQ ID NO:38.
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
XX osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
XX antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
XX immune disease; inflammation; blood disorder; tumour; ss.
XX Homo sapiens.
XX WO200006698-A1.
XX 10-FEB-2000.
XX 29-JUL-1999; 99WO-US17130.
XX 30-JUL-1998; 98US-0094657.
XX 05-AUG-1998; 98US-0095486.
XX 06-AUG-1998; 98US-0095454.
XX 06-AUG-1998; 98US-0095455.
XX 12-AUG-1998; 98US-0096319.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
XX WPI; 2000-195282/17.
XX P-PSDB; AA91478.
XX New isolated human genes and the secreted polypeptides they encode,
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders -
XX Claim 1; Page 392; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
XX human secreted proteins given in AA91451 to AA91691. The human secreted
XX proteins can have activities based on the tissues and cells they are
XX expressed in. Examples of the activities are: cytostatic;
XX immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
XX antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
XX antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
XX corresponding secreted proteins are useful for preventing, treating or
XX ameliorating medical conditions, e.g. by protein or gene therapy. Also
XX pathological conditions can be diagnosed by determining the amount of the
XX proteins in a sample or by determining the presence of mutations in the
XX polynucleotides. Specific uses are described for each of the
XX polynucleotides, based on which tissues they are most highly expressed
XX in, and include developing products for the diagnosis or treatment of
XX cancer, tumours, neurodegenerative disorders, developmental abnormalities
XX and foetal deficiencies, blood disorders, diseases of the immune system,
XX autoimmune diseases, hepatic and renal disease, inflammation,
XX allergies, Alzheimer's and behavioural disorders, schizophrenia,
XX osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
XX transplant rejection, diabetes, asthma, sepsis, acne, psoriasis.
```

CC cardiovascular disorders, reproductive disorders, gastrointestinal  
CC disorders, respiratory disorders and metabolic disorders. The proteins  
CC or polynucleotides can also be used as food additives or preservatives.  
CC The proteins are also useful for identifying their binding partners.  
CC AAA26337 to AAA26345 and AA91450 are sequences used in the  
CC exemplification of the present invention.  
XX  
XX  
SQ Sequence 1048 BP; 372 A; 203 C; 182 G; 291 T; 0 other;  
  
Query Match 94.1%; Score 102.6; DB 21; Length 1048;  
Best Local Similarity 96.3%; Pred. No. 0.00052;  
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 aaataaaatttcaa 60  
Dy 932 aa 991  
  
Qy 61 aaagg 109  
Dy 992 aaagg 1040  
  
RESULT 12  
AAZ91918  
ID AAZ91918 standard; CDNA; 1051 BP.  
XX  
XX AAZ91918;  
XX  
DT 08-JUN-2000 (first entry)  
XX  
DE Murine mahogany protein coding sequence akml004.  
XX  
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
OS Mus sp.  
XX  
PN WO200005373-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 21-JUL-1999; 99WO-US16484.  
XX  
PR 21-JUL-1998; 98US-0093630.  
PR 20-OCT-1998; 98US-0104978.  
PR 05-FEB-1999; 99US-0245041.  
XX  
XX (MILL-) MILLENIUM PHARM INC.  
XX  
PA  
PI Moore K, Nagle DL;  
XX  
XX WPI: 2000-195103/17.  
DR P-PSDB; AAY81805.  
XX  
XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
PT treatment of body weight disorders.  
PT  
PS Claim 1; Fig 9a; 188pp; English.  
XX  
XX This sequence represents a murine mahogany gene of the invention.  
CC The mahogany genes are used: (i) to produce recombinant mahogany (mg)  
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
CC therapeutics; (iii) as a source of diagnostic probes and primers for  
CC detecting expression of mg genes or mutations, regulatory defects, in  
CC this gene, or for isolation of related sequences; and  
CC (iv) in (cell-based) gene therapy. (II) are used to raise specific  
CC antibodies (Ab); to identify other (extra)cellular products involved in  
CC weight regulation, and to screen for agents that disrupt interaction  
CC between (II) and other macromolecules. The Ab are used to detect abnormal  
CC levels (or function) of (II) (for diagnosis, prognosis or monitoring of

CC treatment); to evaluate (II)-expressing cells intended for cell therapy,  
CC and as therapeutic mg inhibitors. Cells that express the mg gene (or  
CC contain the mg polypeptide) are used to identify agents (A) that modulate  
CC mg activity. (A) are potentially useful for the treatment of body weight  
CC disorders, particularly obesity, cachexia or anorexia, or other  
CC conditions associated with the mg gene such as hyperpigmentation,  
CC hyperphagia and disorders that result in increased metabolic rate.  
XX  
XX Sequence 1051 BP; 306 A; 243 C; 313 G; 189 T; 0 other;  
SQ  
  
Query Match 94.1%; Score 102.6; DB 21; Length 1051;  
Best Local Similarity 96.3%; Pred. No. 0.00052;  
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 aaatacaaaatttcaaa 60  
Dy 32 aa 91  
  
Qy 61 aaagg 109  
Dy 92 aaagg 140  
  
RESULT 13  
AAC59409  
ID AAC59409 standard; CDNA; 1126 BP.  
XX  
XX AAC59409;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human secreted protein CDNA #18.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW anti-allergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200056765-A1.  
XX  
PD 28-SEP-2000.  
XX  
PF 16-MAR-2000; 2000WO-US06823.  
XX  
PR 19-MAR-1999; 99US-0125364.  
PR 08-DEC-1999; 99US-0169623.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI: 2000-602215/57.  
DR P-PSDB; AAB33979.  
XX  
XX Nucleic acid molecules encoding human secreted proteins, used in  
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
PT Parkinson's diseases and cancers -  
PS  
PS Claim 1; Page 337; 410pp; English.  
XX  
XX The invention relates to the isolation of genes AAC59392-C59439 encoding  
CC 48 human secreted proteins AAB33963-B34006. The genes can be used to  
CC generate fusion proteins by linking to the gene for the human  
CC immunoglobulin G Fc portion (SEQID) for increasing the stability of  
CC the fusion protein as compared to the human protein only. The genes and  
CC proteins are useful for preventing, ameliorating or treating medical  
CC conditions, e.g. by protein or gene therapy. The genes are isolated  
CC from a range of human tissues disclosed in the specification. The  
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in

CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, and ulcerative  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d)  
 CC wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 XX  
 SQ Sequence 1126 BP; 368 A; 291 C; 143 G; 324 T; 0 other;

Query Match 94.1%; Score 102.6; DB 21; Length 1126;  
 Best Local Similarity 96.3%; Pred. No. 0.00051;  
 Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 aaataaaatttataaaaaaa 60  
 Db 986 aaaaaaa 1045  
 QY 61 aaaaaaa 109  
 Db 1046 aaaaaaa 1094

RESULT 14  
 AAF98699  
 ID AAF98699 standard; DNA; 1164 BP.  
 XX  
 AC AAF98699;  
 XX  
 DT '02-JUL-2001 (first entry)  
 DE Human ovarian cancer cell expressed sequence 10799.  
 XX  
 XX Human; ovarian cancer; identification; detection; characterisation;  
 KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.  
 KW  
 XX Homo sapiens.  
 OS  
 PN WO200118542-A2.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 01-SEP-2000; 2000WO-US24199.  
 XX  
 PR 03-SEP-1999; 99US-0152547.  
 PR 16-MAR-2000; 2000US-0190347.  
 PR 21-MAR-2000; 2000US-0191321.  
 PR 31-MAY-2000; 2000US-0208382.  
 PR 20-JUL-2000; 2000US-0220467.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 XX  
 PI Lee J, Thomsho P, Lillie J;  
 XX  
 DR WPI; 2001-211428/21.  
 XX  
 PT Detection, assessment, prevention and therapy of ovarian cancer,  
 PT comprises detecting changes in the expression of a variety of markers -  
 XX  
 PS Claim 1; Page 1002; 1198pp; English.

XX The present invention describes a method for assessing whether a patient  
 CC is afflicted with ovarian cancer by comparing: (1) the expression of a  
 CC marker (1) (see AAF98594 to AAF98730), in a patient sample; and (2) the  
 CC normal level of expression of (1) in a control non-ovarian cancer  
 CC sample, where a significant difference between the level of expression  
 CC in (a) and (b) is an indication that the patient is afflicted with  
 CC ovarian cancer. (1) have cytostatic activities and can be used in  
 CC antisense gene therapy. The method, compositions and kits from the

CC present invention can be used for: (1) assessing and treating ovarian  
 CC cancer; (2) making isolated hybridoma, which produces an antibody useful  
 CC for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a  
 CC patient. AAF98573 to AAF98593 represent human kinase marker primers and  
 CC probes which are used in the exemplification of the present invention.  
 XX  
 SQ Sequence 1164 BP; 420 A; 176 C; 195 G; 373 T; 0 other;

Query Match 94.1%; Score 102.6; DB 22; Length 1164;  
 Best Local Similarity 96.3%; Pred. No. 0.0005;  
 Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 aaataaaatttataaaaaaa 60  
 Db 1048 aaaaaaa 1107  
 QY 61 aaaaaaa 109  
 Db 1108 aaaaaaa 1156

RESULT 15  
 AAC99826  
 ID AAC99826 standard; cDNA; 1254 BP.  
 XX  
 AC AAC99826;  
 XX  
 DT 13-MAR-2001 (first entry)  
 DE Human secreted protein gene 9 SEQ ID NO:19.  
 XX  
 XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;  
 KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;  
 KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;  
 KW fungicide; ophthalmological; gene therapy; pathological condition;  
 KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;  
 KW neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;  
 KW cerebrovascular disorder; angiogenesis; nervous system disorder;  
 KW Alzheimer's disease; infection; ocular disorder; corneal infection;  
 KW wound healing; skin aging; food additive; preservative; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200070042-A1.  
 XX  
 PD 23-NOV-2000.  
 XX  
 PF 11-MAY-2000; 2000WO-US12788.  
 XX  
 PR 13-MAY-1999; 99US-0134068.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CB;  
 PI Duan RD, Florence KA, Soppet DR;  
 XX  
 DR WPI; 2000-679828/66.  
 DR P-PSDB; AAB56085.  
 XX  
 PT Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 PS Claim 1; Page 860-861; 1065pp; English.

XX The polynucleotide sequences given in AAC99818 to AAC99977 encode the  
 CC human secreted proteins given in AAB56077 to AAB56362. Human secreted  
 CC proteins have activities based on the tissues and cells the genes are  
 CC expressed in. Examples of activities include: immunosuppressive;  
 CC antirheumatic; antiproliferative; antiproliferative; cardiant;  
 CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;  
 CC virucide; fungicide; and ophthalmological. The human secreted  
 CC polynucleotides and proteins can be used to prevent, treat or ameliorate

CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
CC dogs, chickens or sheep. They are also used in diagnosing a pathological  
CC condition or susceptibility to a pathological condition. Disorders which  
CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid  
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or  
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular  
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders  
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
CC fungi and ocular disorders e.g. corneal infection. The proteins can also  
CC be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The proteins can also be used as a  
CC food additive or preservative to increase or decrease storage  
CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used  
CC in the exemplification of the present invention.

XX  
SQ Sequence 1254 BP; 348 A; 383 C; 323 G; 197 T; 3 other;

Query Match 94.1%; Score 102.6; DB 21; Length 1254;  
Best Local Similarity 96.3%; Pred. No. 0.00049;  
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 aaataaaatttataa 60  
Db 1138 aaaaaa aa  
QY 61 aaagg 109  
Db 1198 aaagg 1246

Search completed: January 3, 2002, 08:59:01  
Job time: 124 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 09:21:52 ; Search time 38.35 Seconds  
(without alignments)  
643.705 Million cell updates/sec

Title: US-09-594-596-1  
Perfect score: 109  
Sequence: 1 aaataaaatttaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

\*Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	97	89.0	144	1	US-08-702-344-26
2	97	89.0	1117	4	US-09-247-373B-33
3	96	88.1	1051	4	US-09-245-041-10
4	96	88.1	1474	4	US-08-821-994-64
5	96	88.1	1733	3	US-09-073-569-64
6	96	88.1	9589	1	US-07-925-695-1
7	96	88.1	9589	1	US-07-925-695-1
8	95	87.2	140	1	US-08-628-417-5
9	95	87.2	240	1	US-08-628-417-5
10	95	87.2	1066	1	US-08-157-101A-4
11	95	87.2	1582	3	US-08-545-196B-10
12	95	87.2	1582	3	US-08-545-196B-12
13	94	86.2	98	1	US-08-088-658-42
14	94	86.2	98	2	US-08-471-907A-42
15	94	86.2	111	4	US-09-297-535-23
16	94	86.2	117	1	US-08-702-344-3
17	94	86.2	121	4	US-09-297-535-20
18	94	86.2	1813	4	US-09-071-224-3
19	94	86.2	1882	4	US-09-370-253-1
20	94	86.2	2082	2	US-08-785-310A-2
21	94	86.2	2184	4	US-08-955-918C-1
22	94	86.2	2447	2	US-09-014-969-14
23	94	86.2	6671	1	US-08-280-443-1
24	94	86.2	6671	1	US-08-457-459-1
25	94	86.2	6671	1	US-08-555-678-1
26	94	86.2	6671	5	PCT-US95-02275-1
27	92	84.4	2671	6	5168051-9

28 91 83.5 1411 4 US-08-964-127-5 Sequence 5, Appli  
29 90 82.6 1641 1 US-08-300-903A-8 Sequence 8, Appli  
30 89 81.7 1098 3 US-09-248-335-35 Sequence 35, Appli  
31 88 80.7 2852 3 US-09-027-137-2 Sequence 2, Appli  
32 87 79.8 1198 3 US-09-248-335-27 Sequence 27, Appli  
33 87 79.8 1817 1 US-08-473-981A-5 Sequence 5, Appli  
34 87 79.8 1817 2 US-08-474-087-5 Sequence 5, Appli  
35 86 78.9 2239 4 US-09-196-390-1 Sequence 1, Appli  
36 84 77.1 630 1 US-08-185-414E-1 Sequence 1, Appli  
37 83 76.1 1602 3 US-08-530-950-3 Sequence 3, Appli  
38 83 76.1 1602 3 US-08-888-429A-3 Sequence 3, Appli  
39 83 76.1 1602 4 US-09-149-879-3 Sequence 3, Appli  
40 80 73.4 635 1 US-08-453-633A-35 Sequence 35, Appli  
41 80 73.4 635 1 US-08-416-336-5 Sequence 5, Appli  
42 80 73.4 635 2 US-08-456-460C-35 Sequence 35, Appli  
43 80 73.4 635 5 PCT-US94-05354-35 Sequence 35, Appli  
44 80 73.4 1307 2 US-08-960-022-17 Sequence 17, Appli  
45 80 73.4 1736 3 US-09-182-816-22 Sequence 22, Appli

#### ALIGNMENTS

RESULT 1  
US-08-702-344-26  
; Sequence 26, Application US/08702344  
; Patent No. 5723315  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John  
; APPLICANT: Lavallie, Edward  
; APPLICANT: Racie, Lisa  
; APPLICANT: Merberg, David  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Spaulding, Vikki  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/702,344  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brown, Scott A.  
; REGISTRATION NUMBER: 32,724  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 498-8224  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 144 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; US-08-702-344-26

Query Match 89.0%; Score 97; DB 1; Length 144;  
Best Local Similarity 100.0%; Pred. No. 2,1e-22;  
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







```

RESULT 15
US-09-297-535-23
; Sequence 23, Application US/09297535
; Patent No. 6268199
; GENERAL INFORMATION:
; APPLICANT: Meulenber, Johanna J.M.
; APPLICANT: Poi, Johannes M.A.
; APPLICANT: Bos-de Ruijter, Judy N.A.
; TITLE OF INVENTION: Infectious clones of RNA viruses and vaccines and
; FILE OF INVENTION: diagnostic assays derived thereof.
; FILE REFERENCE: P20749S00
; CURRENT APPLICATION NUMBER: US/09/297,535
; CURRENT FILING DATE: 1999-10-12
; EARLIER APPLICATION NUMBER: PCT/NL97/00593
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: 96203024.3
; EARLIER FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 23
; LENGTH: 111
; TYPE: RNA
; ORGANISM: Lelystad virus
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: (1)..(111)
; OTHER INFORMATION: /note="3' end"
US-09-297-535-23

```

```

Query Match      86.2%; Score 94; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. NO. 1.8e-21;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db * 1 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 74 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 94
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: January 3, 2002, 10:25:57  
Job time: 3845 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 1340.09 Seconds  
(without alignments)  
1341.845 Million cell updates/sec

Title: US-09-594-596-1

Perfect score: 109

Sequence: 1 aaataaaattttaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb\_ba:\*

2: gb\_btg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_om:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_sy:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_hgo\_hum:\*

31: em\_hgo\_inv:\*

32: em\_hgo\_rod:\*

33: em\_hgt\_hum:\*

34: em\_hgt\_inv:\*

35: em\_hgt\_rod:\*

36: em\_hgt\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Result No.	Score	Query Match	Length	DB	ID	Description
1	105.4	96.7	1801	14	SCU78817	U78817 Saccharomyc
2	103.4	94.9	2506	9	AK026927	AK026927 Homo sapi
3	103.4	94.9	3060	6	AX055476	AX055476 Sequence
4	103.4	94.9	3060	6	AX089948	AX089948 Sequence
5	102.8	94.3	596	9	BC008387	BC008387 Homo sapi
6	102.6	94.1	396	6	AX093215	AX093215 Sequence
7	102.6	94.1	1164	6	AX098194	AX098194 Sequence
8	102.6	94.1	1589	9	HS0801509	HS0801509 Homo sapi
9	102.6	94.1	9589	9	E07361	E07361 grNA of Hep
10	102.6	94.1	9589	6	E07362	E07362 cDNA of Hep
11	102.6	94.1	9589	6	I12861	I12861 Sequence 2
12	102.6	94.1	214139	14	HPCPOLP	D00944 Hepatitis C
13	102.6	94.1	214139	2	AC022129	AC022129 Homo sapi
14	102.2	93.8	161	6	A08916	A08916 H.sapiens (
15	102.2	93.8	353	11	CNS08EJ3	AL395173 T7 end of
16	102.2	93.8	951	9	HS0801406	AL133560 Homo sapi
17	102.2	93.8	2417	9	HS0800967	AL117457 Homo sapi
18	102.2	93.8	2773	6	AX092302	AX092302 Sequence
19	101.8	93.4	1843	9	BC009470	BC009470 Homo sapi
20	101.6	93.2	837	9	S78214	S78214 APC-tumor s
21	101.6	93.2	872	6	AX068322	AX068322 Sequence
22	101.6	93.2	1243	9	HS0802621	AL162083 Homo sapi
23	101.6	93.2	1489	9	AB048953	AB048953 Macaca fa
24	101.6	93.2	1892	9	HS0801812	AL136844 Homo sapi
25	101.6	93.2	2118	9	HS0800970	AL117460 Homo sapi
26	101.6	93.2	2296	9	HS0802712	AL1359601 Homo sapi
27	101.6	93.2	2621	10	MUSBRED	L31356 Mus musculu
28	101.6	93.2	3924	10	MUSBRED	L31357 Mus musculu
29	101.6	93.2	5379	9	HS0801286	AL122123 Homo sapi
30	101.6	93.2	6644	6	E23356	E23356 Virus vecto
31	101.6	93.2	7372	6	E23357	E23357 Virus vecto
32	101.6	93.2	7797	6	E23355	E23355 Virus vecto
33	101.6	93.2	7996	6	E23359	E23359 Virus vecto
34	101.6	93.2	74896	2	AC093310	AC093310 Homo sapi
35	101.6	93.2	169170	2	AC093281	AC093281 Homo sapi
36	101.6	93.2	223538	2	AC087150	AC087150 Mus muscu
37	101.6	93.2	256673	2	AC087146	AC087146 Mus muscu
38	101.4	93.0	396	6	AX093235	AX093235 Sequence
39	101.2	92.8	381	3	AF159974	AF159974 Butchus ma
40	101.2	92.8	2785	9	HS0800418	AL050108 Homo sapi
41	101.2	92.8	45685	2	AC087168	AC087168 Homo sapi
42	101	92.7	1569	9	HS0801353	AL133080 Homo sapi
43	101	92.7	9580	14	AF054250	AF054250 Hepatitis
44	101	92.7	154272	2	AF001888	AF001888 Homo sapi
45	100.6	92.3	117	6	I89931	I89931 Sequence 3

#### ALIGNMENTS

RESULT 1  
SCU78817 SCU78817 1801 bp RNA 29-JUL-1997  
LOCUS Saccharomycetes cerevisiae killer virus M1, complete genome.  
DEFINITION U78817  
ACCESSION U78817  
VERSION U78817.1 GI:1699029  
KEYWORDS  
SOURCE  
ORGANISM  
Saccharomycetes cerevisiae killer virus M1.  
Saccharomycetes cerevisiae killer virus M1  
Viruses; unclassified viruses.  
REFERENCE  
1 (bases 1 to 1801)  
Russell,P.J., Bennett,A.M., Love,Z. and Baggott,D.M.  
Cloning, sequencing and expression of a full-length cDNA copy of  
the M1 double-stranded RNA virus from the yeast, Saccharomycetes  
cerevisiae  
Yeast 13 (9), 829-836 (1997)  
MEDLINE 9737790  
REFERENCE  
2 (bases 1 to 1801)  
Russell,P.J., Bennett,A.M., Love,Z. and Baggott,D.M.  
Direct Submission  
TITLE

FEATURES		Location/Qualifiers	
source	1..596		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="4052341"		
	/tissue_type="Kidney, hypernephroma"		
	/clone_lib="NIH_MGC_58"		
	/lab_host="DH105"		
	/note="Vector: pDNR-LIB"		
BASE COUNT	286 a 82 c 107 g 121 t		
ORIGIN			
Query Match 94.3%; Score 102.8; DB 9; Length 596;			
Best Local Similarity 98.1%; Pred. NO. 0.009; 2; Indels 0; Gaps 0;			
Matches 104; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	2 aaataaaatttataaaaaa	61	
Db	441 AAATAAAATCAAAAAA	500	
Qy	62 aaaaaa	107	
Db	501 AAAAAA	546	
RESULT 6			
AX093215/c			
LOCUS	AX093215	396 bp	DNA
DEFINITION	Sequence 33 from Patent WO0118046.	PAT	30-MAR-2001
ACCESSION	AX093215		
VERSION	AX093215.1	GI:13509664	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Xu,J. and Stolk,J.A.		
TITLE	Ovarian tumor sequences and methods of use therefor		
JOURNAL	Patent: WO 0118046-A 33 15-MAR-2001;		
	CORIXA CORPORATION (US)		
FEATURES		Location/Qualifiers	
source	1..396		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
misc_feature	1..396		
	/note="n = A,T,C or G"		
BASE COUNT	60 a 17 c 34 g 162 t 123 others		
ORIGIN			
Query Match 94.1%; Score 102.6; DB 6; Length 396;			
Best Local Similarity 96.3%; Pred. NO. 0.01;			
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	1 aaataaaatttataaaaa	60	
Db	109 AAAAAA	50	
Qy	61 aaaaaa	109	
Db	49 AAAAAA	1	
RESULT 7			
AX098194			
LOCUS	AX098194	1164 bp	DNA
DEFINITION	Sequence 106 from Patent WO0118542.	PAT	30-MAR-2001
ACCESSION	AX098194		
VERSION	AX098194.1	GI:13515278	
KEYWORDS	human.		
SOURCE			

```

Qy 1 aaataaaattttaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy 61 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 109
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480

RESULT 10
E07362/c
LOCUS E07362 9589 bp RNA PAT 29-SEP-1997
DEFINITION CDNA of Hepatitis non-A non-B virus.
ACCESSION E07362
VERSION E07362.1 GI:2175501
KEYWORDS JP 1994121689-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 9589)
AUTHORS Okamoto,H. and Nakamura,T.
TITLE NON-A NON-B HEPATITIS VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
ANTIGEN AND ANTIBODY DETECTION SYSTEM
JOURNAL Patent: JP 1994121689-A 2 06-MAY-1994;
NAKAMURA TETSUO
COMMENT OS Hepatitis non-A non-B virus
PN JP 1994121689-A/2
PD 06-MAY-1994
PF 09-AUG-1991 JP 1991287402
PI OKAMOTO HIROAKI, NAKAMURA TETSUO
PC C12N15/51,A61K39/29,A61K39/395,A61K39/395,C07K13/00,C12P21/02,
C12P21/08,
PC G01N33/53,G01N33/576,G01N33/577//A61B10/00;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH Key Location/Qualifiers
FT source 1..9589
FT /strain='HCV-J6'.
FEATURES
source Location/Qualifiers
1..9589
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 1968 a 2820 c 2635 g 2166 t
ORIGIN

Query Match 94.1%; Score 102.6; DB 6; Length 9589;
Best Local Similarity 96.3%; Pred. No. 0.0061;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aaataaaattttaaataaaataaaataaaataaaataaaataaaataaaataaa 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy 61 aaaaaataaaataaaataaaataaaataaaataaaataaaataaaataaaataaa 109
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480

RESULT 12
HPCPOLP/c
LOCUS HPCPOLP 9589 bp RNA VRL 20-JUN-1998
DEFINITION Hepatitis C virus genomic RNA for polyprotein, complete cds.
ACCESSION D00944
VERSION D00944.1 GI:221650
KEYWORDS polyprotein.
SOURCE Hepatitis C virus (isolate:HC-J6) CDNA to genomic RNA.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1 (bases 1 to 9589)
AUTHORS Okamoto,H., Okada,S., Sugiyama,Y., Kurai,K., Iizuka,H., Machida,A.,
Miyakawa,Y. and Mayumi,M.
TITLE Nucleotide sequence of the genomic RNA of hepatitis C virus
isolated from a human carrier: comparison with reported isolates
for conserved and divergent regions
J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
92044440
2 (sites)
Han,J.H. and Houghton,M.
Group specific sequences and conserved secondary structures at the
3' end of HCV genome and its implication for viral replication
Nucleic Acids Res. 20 (13), 3520 (1992)
92335016
3 (sites)
Hotta,H., Doi,H., Hayashi,T., Purwanta,M., Soemarto,W.,
Mizokami,M., Ohba,K. and Homma,M.
Analysis of the core and E1 envelope region sequences of a novel
variant of hepatitis C virus obtained in Indonesia
Arch. Virol. 136 (1-2), 53-62 (1994)
94270990
COMMENT These data kindly submitted in computer readable form by: Hiroaki
Okamoto
Immunology Division
Jichi Medical School
Kawachi-gun
Tochigi-ken 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557
Location/Qualifiers
1..9589
/organism='Hepatitis C virus'
/isolate='HC-J6'
/db_xref='taxon:11103'
341..9442
CDS

```



Qy	1	aaataaaaattttaaaaaaatccc
Db	9588	AA
Qy	61	aaagg
Db	9528	AAAGG
 RESULT 10 E07362/c		
LOCUS	E07362	cdna of Hepatitis non-A non-B virus.
DEFINITION	E07362	PAT 29-SEP-1997
ACCESSION	E07362	
VERSION	E07362.1	GI:2175501
KEYWORDS	JP 1994121689-A/2.	
SOURCE	unidentified.	
ORGANISM	unclassified.	
REFERENCE	1 (bases 1 to 9589)	
AUTHORS	Okamoto,H. and Nakamura,T.	
TITLE	NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,	
JOURNAL	ANTIGEN AND ANTIBODY DETECTION SYSTEM	
COMMENT	Patent: JP 1994121689-A 2 06-MAY-1994;	
	NAKAMURA TETSUO	
	OS Hepatitis non-A non-B virus	
	PN JP 1994121689-A/2	
	PD 06-MAY-1994	
	PF 09-AUG-1991 JP 1991287402	
	PI OKAMOTO HIROAKI, NAKAMURA TETSUO	
	PC C12N15/31,A61K39/395,A61K39/395,C07K13/00,C12P21/02,	
	PC C12P21/08,	
	PC G01N33/53,G01N33/576,G01N33/577//A61B10/00;	
	CC strandedness: Single;	
	CC topology: Linear;	
	CC hypothetical: No;	
	CC anti-sense: No;	
FH Key	Location/Qualifiers	
FT source	1..9589	
FT	/strain='HCV-J6'.	
FEATURES	Location/Qualifiers	
source	1..9589	
	/organism='Hepatitis non-A non-B virus'	
BASE COUNT	1968 a 2820 c 2635 g 2166 t	
ORIGIN		
 Query Match 94.1%; Score 102.6; DB 6; Length 9589; Best Local Similarity 96.3%; Pred. No. 0.0061. Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;		
y	1	aaataaaaattttaaaaaaatccc
b	9588	AAAGG
y	61	aaagg
b	9528	AAAGG
 RESULT 11 I12861/c		
XCU	I12861	Sequence 2 from patent US 5428145.
FINITION	Sequence 2 from patent US 5428145.	
CESSION	I12861	PAT 26-JUL-1995
RSION	I12861.1	GI:910242
YWORDS	.	
URCE	Unknown.	
ORGANISM	Unknown.	







